GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 05:06:07 ; Search time 206.5 Seconds
(without alignments)
261.447 Million cell updates/sec

Title: US-09-927-585A-2
Perfect score: 20
Sequence: 1 tgctaattcggtgctacctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	14	AAQ48541	HPV E6/7 region pr
2	20	100.0	20	14	AAQ48565	HPV E6/7 region pr
3	20	100.0	20	24	AAS18613	Low hazard strain
4	20	100.0	731	10	AAN91601	Partial nucleotide
5	20	100.0	1107	18	AAT59430	Human papillomavir
6	20	100.0	1128	18	AAT59431	Human papillomavir
7	20	100.0	1398	18	AAT59432	Human papillomavir
8	20	100.0	4770	18	AAT59450	Human papillomavir
9	20	100.0	8010	17	AAT13413	Human papillomavir
c 10	18	90.0	23	25	ABX12328	Fluorescent multip
11	15.8	79.0	281	22	AAS29985	Human lung antigen
12	15.8	79.0	281	22	AAS28494	Genomic sequence #
13	15.8	79.0	287	22	AAS29864	Human lung antigen
14	15.8	79.0	287	22	AAS28028	Novel cDNA encodin
c 15	15.8	79.0	907	25	ABZ52394	Aspergillus oryzae
c 16	15.8	79.0	4682	25	ABT17899	Aspergillus fumiga
17	15.4	77.0	640	23	ABV45803	Human prostate exp
18	15.4	77.0	650	23	ABV16004	Human prostate exp
c 19	15.4	77.0	1111	23	AAS64558	DNA encoding novel
c 20	15.4	77.0	1368	24	ABN68579	Streptococcus poly
c 21	15.4	77.0	1368	24	ABN70478	Streptococcus poly
22	15.4	77.0	2564	23	ABV30275	Human prostate exp
c 23	15.4	77.0	3340	21	AAA79267	Pinus radiata cell
c 24	15.4	77.0	2155561	24	ABN71527	Streptococcus poly
25	15.2	76.0	20	14	AAQ48545	HPV E6/7 region pr
26	15.2	76.0	245	25	ABX63752	Human cDNA #752 di
27	15.2	76.0	720	10	AAN91603	Partial nucleotide
28	15.2	76.0	767	24	AAL48310	HPV11-E6/E7 region
c 29	15.2	76.0	1715	22	ABA18187	Human nervous syst
c 30	15.2	76.0	1716	22	ABA18189	Human nervous syst
31	15.2	76.0	12587	22	AAK87087	Human immune/haema
32	15.2	76.0	65921	21	AAZ89046	Human nibrin DNA.
c 33	15.2	76.0	349980	22	AAF86431	Pyrococcus abyssi
c 34	14.8	74.0	94	21	AAC23951	Human secreted pro
35	14.8	74.0	1285	19	AAV59580	Human secreted pro
36	14.8	74.0	1285	24	ABS73567	Human cDNA #1 for
c 37	14.8	74.0	1299	23	ABL18657	Drosophila melanog
c 38	14.8	74.0	2294	23	ABL25274	Drosophila melanog
39	14.8	74.0	2537	22	AAH17687	Human cDNA sequenc
40	14.8	74.0	3330	23	ABL21563	Drosophila melanog
41	14.8	74.0	3363	23	ABL18656	Drosophila melanog
42	14.8	74.0	4165	23	ABL05520	Drosophila melanog
43	14.8	74.0	4520	23	ABL05532	Drosophila melanog
44	14.8	74.0	4554	23	ABL10638	Drosophila melanog
c 45	14.8	74.0	5626	23	ABL21562	Drosophila melanog

ALIGNMENTS

RESULT 1
AAQ48541
ID AAQ48541 standard; DNA; 20 BP.
XX
AC AAQ48541;
XX
DT 22-FEB-1994 (first entry)
XX
DE HPV E6/7 region probe.
XX
KW Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.
XX
OS Synthetic.
XX
PN JP05192200-A.
XX
PD 03-AUG-1993.
XX
PF 19-AUG-1991; 91JP-0230839.
XX
PR 20-AUG-1990; 90JP-0217067.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
DR WPI; 1993-277497/35.
XX
PT Detecting benign and/or malignant human papilloma virus - by
PT detecting DNA sequence of E6 and/or E7 region of human papilloma
PT virus
XX
PS Disclosure; Page 13; 18pp; Japanese.

```
XX
CC   The probe is used to detect benign and/or malignant human papilloma
CC   virus. The probe binds to the E6 and/or E7 region of the virus.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 2
AAQ48565
ID   AAQ48565 standard; DNA; 20 BP.
XX
AC   AAQ48565;
XX
DT   22-FEB-1994  (first entry)
XX
DE   HPV E6/7 region probe.
XX
KW   Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.
XX
OS   Synthetic.
XX
PN   JP05192200-A.
XX
PD   03-AUG-1993.
XX
PF   19-AUG-1991;   91JP-0230839.
XX
PR   20-AUG-1990;   90JP-0217067.
XX
PA   (TAKI ) TAKARA SHUZO CO LTD.
XX
DR   WPI; 1993-277497/35.
XX
PT   Detecting benign and/or malignant human papilloma virus - by
PT   detecting DNA sequence of E6 and/or E7 region of human papilloma
PT   virus
XX
PS   Disclosure; Page 15; 18pp; Japanese.
XX
CC   The probe is used to detect benign and/or malignant human papilloma
CC   virus. The probe binds to the E6 and/or E7 region of the virus.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 3
AAS18613
ID   AAS18613 standard; DNA; 20 BP.
XX
AC   AAS18613;
XX
DT   26-FEB-2002  (first entry)
XX
DE   Low hazard strain human papillomavirus (HPV) E6, PCR primer.
XX
KW   Human papillomavirus; HPV; early gene; E6; PCR primer;
KW   low hazard strain; ss.
XX
OS   Human papillomavirus.
XX
PN   WO200185994-A2.
XX
PD   15-NOV-2001.
XX
PF   02-MAY-2001; 2001WO-IB00771.
XX
PR   05-MAY-2000; 2000IT-MO00091.
XX
PA   (BIOA-) BIOANALISI CENT SUD DI PERSEU SINIBLADO.
XX
PI   Perseu S,  De Montis A,  Floris MM;
XX
DR   WPI; 2002-075253/10.
XX
PT   Identifying viral DNA of human papilloma virus in cellular material
PT   collected from an individual, by using polymerase and identifying
PT   sequences complementary to regions E of the viral genome using a
PT   reagent -
XX
PS   Claim 11; Page 12; 17pp; English.
XX
CC   The invention describes a novel method of identifying viral DNA of human
CC   papillomavirus (HPV) in cellular material collected from an individual.
CC   The method comprises extracting DNA from the cellular material,
CC   amplifying the DNA by using a polymerase, and identifying sequences
CC   complementary to sequences in the regions E (early genes) e.g. E6 and E7
CC   of the HPV genome using an identifying reagent. The presence of HPV in
CC   cells is now based on identification of a nucleotide sequence in the L1
CC   region of the genome. Identification of E6, E7 and L1 make it possible to
CC   identify the presence of HPV virus in every stage of infection and to
CC   distinguish between virus strains having low hazard and virus strains
CC   having middle/high hazard. This is the primer used to identify the E6
CC   region of the low hazard HPV strains in infected cells, described in the
CC   method of the invention.
XX
SQ   Sequence 20 BP; 3 A; 5 C; 5 G; 7 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 24;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 4
AAN91601
ID   AAN91601 standard; DNA; 731 BP.
XX
AC   AAN91601;
XX
DT   17-JUL-1990  (first entry)
XX
DE   Partial nucleotide sequence (5' end) of human papilloma virus (HPV)
DE   type 6 (HPV-6).
XX
KW   Human papilloma virus; type 6; in situ hybridisation assay;
KW   cellular smear; benign cervical wart; cervical cancer.
XX
OS   Human papilloma virus.
XX
PN   WO8902934-A.
XX
PD   06-APR-1989.
XX
PF   30-SEP-1988;   88WO-US03367.
XX
PR   02-OCT-1987;   87US-0103979.
XX
```

PA (MICR-) MICROPROBE CORP.
XX
PI Schwartz DE, Adams TH;
XX
DR WPI; 1989-114406/15.
XX
PT Hybridisation test for human papilloma virus in cell smears -
PT by reaction with long labelled probe specific for particular
PT virus types, esp. for examining cervical smears
XX
PS Disclosure; ; 39pp; English.
XX
CC The patent is for a rapid in situ hybridisation assay for detecting and
CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to
CC a support in absence of aldehyde-based crosslinking reagents. The assay
CC comprises: (1) combining nucleic acid in the sample with at least one
CC detectable probe able to hybridise with 1 or more HPV types; and (2)
CC detecting presence or absence of hybrid complexes. Opt. several probes
CC are used, eg one for HPV types 6 and 11, associated with benign warts,
CC and one for types 16, 18, 31, 33 and 35, associated with cervical
CC cancer. The assay can differentiate between HPV types. It is esp. used
CC as a secondary test. The probes can be synthesised or cloned.
XX
SQ Sequence 731 BP; 245 A; 147 C; 159 G; 180 T; 0 other;

Query Match 100.0%; Score 20; DB 10; Length 731;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTAATTCGGTGCTACCTG 20
||||||||||||||||||||
Db 400 TGCTAATTCGGTGCTACCTG 419

RESULT 5
AAT59430
ID AAT59430 standard; DNA; 1107 BP.
XX
AC AAT59430;
XX
DT 26-MAY-1997 (first entry)
XX
DE Human papillomavirus 6b E6/E5a/E4 DNA cassette CSL690.
XX
KW HPV; polyprotein; vaccine; wart; condylomata acuminata;
KW cervical cancer; ss.
XX
OS Human papillomavirus type 6b.
XX
PN WO9705164-A1.
XX
PD 13-FEB-1997.
XX
PF 26-JUL-1996; 96WO-AU00473.
XX
PR 27-JUL-1995; 95AU-0004439.
XX
PA (CSLC-) CSL LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;
PI Moloney MBH, Webb EA, Williams MP;
XX
DR WPI; 1997-145619/13.
DR P-PSDB; AAW01808.
XX
PT Papilloma virus early open reading frame polyprotein constructs -
PT useful for treating, e.g. genital warts and cervical cancer
XX
PS Example 1; Page 22-23; 62pp; English.
XX
CC DNA construct CSL690 (AAT59430) codes for an E6/E5a/E4 polyprotein
CC (AAW01808), plus C-terminal hexahistidine tag, of human
CC papillomavirus 6b (HPV6b). It was obtd. by separate PCR
CC amplification of the E6, E4 and E5a genes (see also AAT59412-17) of
CC HCV6b and their cloning into vector pSP70. The resulting
CC E6/E5a/E4 cassette was incorporated into pGEX-STOP to allow prodn.
CC of the HPV6b polyprotein in E. coli. The E6/E5a/E4 DNA construct,
CC or the encoded polypeptide, can be used as a vaccine to generate a
CC humoral and/or cellular immune response against HPV, esp. for
CC treating genital warts and cervical cancer.
XX
SQ Sequence 1107 BP; 337 A; 268 C; 213 G; 289 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 1107;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGCTAATTCGGTGCTACCTG 20
||||||||||||||||||||
Db 299 TGCTAATTCGGTGCTACCTG 318

RESULT 6
AAT59431
ID AAT59431 standard; DNA; 1128 BP.
XX
AC AAT59431;
XX
DT 26-MAY-1997 (first entry)
XX
DE Human papillomavirus 6b E6/E7/E4 DNA cassette CSL760.
XX
KW HPV; polyprotein; vaccine; wart; condylomata acuminata;
KW cervical cancer; ss.
XX
OS Human papillomavirus type 6b.
XX
PN WO9705164-A1.
XX
PD 13-FEB-1997.
XX
PF 26-JUL-1996; 96WO-AU00473.
XX
PR 27-JUL-1995; 95AU-0004439.
XX
PA (CSLC-) CSL LTD.
PA (UYQU) UNIV QUEENSLAND.
XX
PI Cox JC, Edwards SJ, Frazer I, Margetts MB, McMillan NAJ;
PI Moloney MBH, Webb EA, Williams MP;
XX
DR WPI; 1997-145619/13.
DR P-PSDB; AAW01809.
XX
PT Papilloma virus early open reading frame polyprotein constructs -
PT useful for treating, e.g. genital warts and cervical cancer
XX
PS Example 1; Page 24-25; 62pp; English.
XX
CC DNA construct CSL760 (AAT59431) codes for an E6/E7/E4 polyprotein
CC (AAW01809) of human papillomavirus 6b (HPV6b), plus a C-terminal
CC hexahistidine tag. It was obtd. by separate PCR amplification of
CC the E6, E4 and E7 genes (see also AAT59412-15 and AAT59418-19) of
CC HCV6b and their cloning into vector pSP70. The resulting
CC E6/E7/E4 cassette was incorporated into pGEX-STOP to allow prodn.
CC of the HPV6b polyprotein in E. coli. The E6/E7/E4 DNA construct,
CC or the encoded polypeptide, can be used as a vaccine to generate a
CC humoral and/or cellular immune response against HPV, esp. for
CC treating genital warts and cervical cancer.
XX
SQ Sequence 1128 BP; 359 A; 274 C; 240 G; 255 T; 0 other;

Query Match 100.0%; Score 20; DB 18; Length 1128;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 7
AAT59432
ID   AAT59432 standard; DNA; 1398 BP.
XX
AC   AAT59432;
XX
DT   26-MAY-1997  (first entry)
XX
DE   Human papillomavirus 6b E6/E7/E5a/E4 DNA cassette CSL673.
XX
KW   HPV; polyprotein; vaccine; wart; condylomata acuminata;
KW   cervical cancer; ss.
XX
OS   Human papillomavirus type 6b.
XX
PN   WO9705164-A1.
XX
PD   13-FEB-1997.
XX
PF   26-JUL-1996;   96WO-AU00473.
XX
PR   27-JUL-1995;   95AU-0004439.
XX
PA   (CSLC-) CSL LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Cox JC,  Edwards SJ,  Frazer I,  Margetts MB,  McMillan NAJ;
PI   Moloney MBH, Webb EA,  Williams MP;
XX
DR   WPI; 1997-145619/13.
DR   P-PSDB; AAW01810.
XX
PT   Papilloma virus early open reading frame polyprotein constructs -
PT   useful for treating, e.g. genital warts and cervical cancer
XX
PS   Example 1; Page 26-28; 62pp; English.
XX
CC   DNA construct CSL673 (AAT59432) codes for an E6/E7/E5a/E4 polyprotein
CC   (AAW01810) of human papillomavirus 6b (HPV6b), plus a C-terminal
CC   hexahistidine tag.  It was obtd. by separate PCR amplification of
CC   the E6, E4, E5a and E7 genes (see also AAT59412-19 of HCV6b and their
CC   cloning into vector pSP70.  The resulting E6/E7/E5a/E4 cassette
CC   was incorporated into pGEX-STOP to allow prodn. of the HPV6b
CC   polyprotein in E. coli.  The E6/E7/E5a/E4 DNA construct, or the
CC   encoded polypeptide, can be used as a vaccine to generate a
CC   humoral and/or cellular immune response against HPV, esp. for
CC   treating genital warts and cervical cancer.
XX
SQ   Sequence 1398 BP; 429 A; 330 C; 283 G; 356 T; 0 other;

  Query Match           100.0%;  Score 20;  DB 18;  Length 1398;
  Best Local Similarity  100.0%;  Pred. No. 0.48;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 8
AAT59450
ID   AAT59450 standard; DNA; 4770 BP.
XX
AC   AAT59450;
XX
DT   31-MAY-1997  (first entry)
XX
DE   Human papillomavirus 6 polyprotein DNA.
XX
KW   HPV; polyprotein; vaccine; wart; condylomata acuminata;
KW   cervical cancer; ds.
XX
OS   Human papillomavirus type 6.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..4764
FT                   /*tag= a
FT                   /product= hexaHis-HPV6 polyprotein
FT   misc_RNA        1..108
FT                   /*tag= b
FT                   /function= Tag used for protein purification
FT                   /note= "hexaHis leader sequence form pTrcHisA"
FT   mRNA            115..1218
FT                   /*tag= c
FT                   /note= "HPV-6 E2"
FT   mRNA            1225..1551
FT                   /*tag= d
FT                   /note= "HPV-6 E4"
FT   mRNA            1558..1830
FT                   /*tag= e
FT                   /note= "HPV-6 E5a"
FT   mRNA            1837..2052
FT                   /*tag= f
FT                   /note= "HPV-6 E5b"
FT   mRNA            2059..2508
FT                   /*tag= g
FT                   /note= "HPV-6 E6"
FT   mRNA            2515..2808
FT                   /*tag= h
FT                   /note= "HPV-6 E7"
FT   mRNA            2815..4764
FT                   /*tag= i
FT                   /note= "HPV-6 E1"
XX
PN   WO9705164-A1.
XX
PD   13-FEB-1997.
XX
PF   26-JUL-1996;   96WO-AU00473.
XX
PR   27-JUL-1995;   95AU-0004439.
XX
PA   (CSLC-) CSL LTD.
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Cox JC,  Edwards SJ,  Frazer I,  Margetts MB,  McMillan NAJ;
PI   Moloney MBH, Webb EA,  Williams MP;
XX
DR   WPI; 1997-145619/13.
DR   P-PSDB; AAW01816.
XX
PT   Papilloma virus early open reading frame polyprotein constructs -
PT   useful for treating, e.g. genital warts and cervical cancer
XX
PS   Example 6; Page 35-40; 62pp; English.
XX
CC   A DNA construct (AAT59450) in vector pTrcHisA codes for an
CC   E2/E4/E5a/E5b/E6/E7/E1 polyprotein (AAW01816) of human papillomavirus
CC   6 (HPV6), plus an N-terminal hexahistidine tag.  It was obtd. by
CC   amplification of the individual early genes (see also AAT59436-49)
CC   and their cloning into pTrcHisA.  The DNA construct, or the encoded
CC   polyprotein, can be used as a vaccine to generate a humoral and/or
CC   cellular immune response against HPV, esp. for treating genital
CC   warts and cervical cancer.
XX
SQ   Sequence 4770 BP; 1597 A; 890 C; 1071 G; 1212 T; 0 other;

  Query Match           100.0%;  Score 20;  DB 18;  Length 4770;
  Best Local Similarity  100.0%;  Pred. No. 0.55;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db       2357 TGCTAATTCGGTGCTACCTG 2376


RESULT 9
AAT13413
ID   AAT13413 standard; cDNA; 8010 BP.
XX
AC   AAT13413;
XX
DT   12-JUN-1996  (first entry)
XX
DE   Human papillomavirus type 6a.
XX
KW   Vaccine; immunotherapy; genetic immunisation; condyloma acuminata;
KW   ss.
XX
OS   Human papillomavirus type 6a.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..555
FT                   /*tag= a
FT                   /label= ORF-E6
FT   CDS             441..827
FT                   /*tag= b
FT                   /label= ORF-E7
FT   CDS             716..2782
FT                   /*tag= c
FT                   /label= ORF-E1
FT   CDS             2695..3830
FT                   /*tag= d
FT                   /label= ORF-E2
FT   CDS             2240..3585
FT                   /*tag= e
FT                   /label= ORF-E4
FT   CDS             3888..4163
FT                   /*tag= f
FT                   /label= ORF-E5
FT   CDS             4379..5803
FT                   /*tag= g
FT                   /label= ORF-L2
FT   CDS             5679..7292
FT                   /*tag= h
FT                   /label= ORF-L1
FT                   /product= major capsid protein
XX
PN   WO9609375-A1.
XX
PD   28-MAR-1996.
XX
PF   18-SEP-1995;   95WO-US11859.
XX
PR   22-SEP-1994;   94US-0310468.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Hofmann KJ,  Jansen KU;
XX
DR   WPI; 1996-188437/19.
XX
PT   DNA encoding human papillomavirus 6a  - used to treat and in
PT   vaccines against human papillomavirus 6a disease, e.g. condyloma
PT   acuminata
XX
PS   Claim 1; Fig 1A-G; 46pp; English.
XX
CC   A DNA molecule (AAT13413) codes for human papillomavirus 6a (HPV6a),
CC   the predominant HPV subtype found in biopsies of condyloma acuminata
CC   (benign lesions of the respiratory and genital mucosa).  It was
CC   isolated from DNA extracted from a vulva condyloma acuminatum lesion
CC   following screening with a probe (see AAT13414) complementary to the 3'
CC   end of the HPV subtype 6b L1 gene.  The DNA can be used to develop
CC   vaccines against HPV6a, to treat conditions caused by HPV6a, for
CC   HPV serotyping, and to produce recombinant HPV6a polypeptides.
XX
SQ   Sequence 8010 BP; 2470 A; 1529 C; 1715 G; 2296 T; 0 other;

  Query Match           100.0%;  Score 20;  DB 17;  Length 8010;
  Best Local Similarity  100.0%;  Pred. No. 0.58;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


RESULT 10
ABX12328/c
ID   ABX12328 standard; DNA; 23 BP.
XX
AC   ABX12328;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Fluorescent multiplex HPV PCR assay primer #4.
XX
KW   Human papillomavirus; HPV; fluorescent multiplex HPV PCR assay; primer;
KW   PCR; ss; polymerase; fluorophore; fluorescence.
XX
OS   Human papillomavirus.
XX
PN   WO2003019143-A2.
XX
PD   06-MAR-2003.
XX
PF   19-AUG-2002; 2002WO-US26964.
XX
PR   23-AUG-2001; 2001US-314383P.
XX
PA   (MERI ) MERCK & CO INC.
XX
PI   Jansen KU,  Taddeo FJ,  Li W,  Dicello AC;
XX
DR   WPI; 2003-312914/30.
XX
PT   Detecting the presence of a human papillomavirus subtype, using
PT   multiple fluorophores, in a nucleic acid-containing sample, useful in
PT   PCR-based assays for identifying HPV subtypes  -
XX
PS   Claim 15; Page 39; 59pp; English.
XX
CC   The invention relates to a method for detecting the presence of a human
CC   papillomavirus (HPV) subtype in a nucleic acid-containing sample. The
CC   method comprises amplifying the nucleic acid in the presence of a
CC   polymerase and a plurality of oligonucleotide sets, allowing digestion of
CC   each fluorescent probe during amplification to dissociate the fluorophore
CC   from the quencher molecule, detecting a change of fluorescence upon
CC   dissociation, and determining if the sample is positive for the HPV
CC   subtype if a change of fluorescence is detected in at least two emission
CC   maxima. Each oligonucleotide set mentioned above consists of a forward
CC   discriminatory PCR primer hybridising to a first location of an HPV
CC   subtype, a reverse discriminatory PCR primer hybridising to a second
CC   location of the HPV subtype downstream of the first location, and a
CC   fluorescent probe labelled with a quencher molecule and a fluorophore
CC   which emits energy at a unique emission maxima, the probe hybridising to
CC   a location of the HPV subtype between the first and the second locations,
CC   where each oligonucleotide set specifically hybridises to a different HPV
CC   amplicon derived from the same HPV subtype. The change of fluorescence
CC   corresponds to the occurrence of nucleic acid amplification. The method
CC   of the present invention is useful in PCR-based assays for detecting HPV
CC   subtypes in clinical samples. This sequence represents a PCR primer used
CC   in the fluorescent multiplex HPV PCR assay of the invention.
XX
SQ   Sequence 23 BP; 8 A; 5 C; 6 G; 4 T; 0 other;
```

```
  Query Match             90.0%;  Score 18;  DB 25;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 3.7;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 CTAATTCGGTGCTACCTG 20
              ||||||||||||||||||
Db         23 CTAATTCGGTGCTACCTG 6


RESULT 11
AAS29985
ID   AAS29985 standard; DNA; 281 BP.
XX
AC   AAS29985;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Human lung antigen genomic DNA #55.
XX
KW   Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW   chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW   cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW   ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW   hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
KW   cerebrovascular disorder; nervous system disorder; bacterial infection;
KW   fungal infection; viral infection; ocular disorder; endocrine disorder;
KW   gastrointestinal disorder; renal disorder; respiratory disorder;
KW   wound healing; skin aging; organ transplantation; food preservative;
KW   tissue regeneration; anti-infertility; food additive.
XX
OS   Homo sapiens.
XX
PN   WO200155303-A2.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01301.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
PR   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
```

```
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-457723/49.
XX
PT   Isolated polypeptide for treating, preventing and/or prognosing
PT   respiratory disorders related to the lung including lung cancers and
PT   also for testing and detection e.g. diagnosis -
XX
PS   Claim 1; SEQ ID No 249; 507pp; English.
XX
CC   Sequences AAS29931-AAS30164 represent genomic DNA molecules, which encode
CC   the lung antigen polypeptides of the invention. Lung antigen polypeptides
CC   and their associated polynucleotides are useful in the diagnosis,
CC   treatment and prevention of various types of disorders in e.g. humans,
CC   mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A
CC   pathological condition can be determined by detecting the presence or
CC   absence of a mutation in a lung antigen polynucleotide. The treatable
CC   disorders include autoimmune diseases such as rheumatoid arthritis,
CC   hyperproliferative disorders such as neoplasms of the breast or liver,
CC   cardiovascular disorders such as cardiac arrest, cerebrovascular
CC   disorders such as cerebral ischaemia, nervous system disorders such as
CC   Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC   ocular disorders such as corneal infection, endocrine disorders such as
CC   premature labour and infertility, gastrointestinal disorders such as
CC   Crohn's disease, renal disorders such as glomerulonephritis and
CC   respiratory disorders such as asthma and pleurisy. The polypeptides can
CC   also be used to aid wound healing, to prevent skin aging due to sunburn,
CC   to maintain organs before transplantation, to regenerate tissues and in
CC   chemotaxis. The polypeptides can also be used as a food additive or
CC   preservative to increase or decrease storage capabilities.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.

  Query Match             79.0%;  Score 15.8;  DB 22;  Length 281;
  Best Local Similarity   89.5%;  Pred. No. 80;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 12
AAS28494
ID   AAS28494 standard; DNA; 281 BP.
XX
AC   AAS28494;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Genomic sequence #334 encoding for novel human respiratory antigen.
XX
KW   Human; respiratory antigen; respiratory disorder; throat disorder;
KW   lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW   anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW   respiratory active; ds.
XX
OS   Homo sapiens.
XX
PN   WO200155448-A1.
XX
PD   02-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01333.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
```

PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476224/51.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory
PT cancers and also for testing and detection e.g. diagnosis -
XX
PS Disclosure; SED ID No 928; 546pp; English.
XX
CC The present invention relates to the isolation of novel human
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful for preventing, treating and/or prognosing
CC disorders related to the respiratory system including throat
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS28161-AAS28764 represent genomic sequences encoding for novel
CC human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 281 BP; 83 A; 63 C; 26 G; 109 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 281;
Best Local Similarity 89.5%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGCTAATTCGGTGCTACCT 19
||||| ||| |||||||||
Db 102 TGCTATTTCTGTGCTACCT 120

RESULT 13
AAS29864
ID AAS29864 standard; cDNA; 287 BP.
XX
AC AAS29864;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human lung antigen cDNA polynucleotide #26.
XX
KW Lung antigen protein; human; mouse; rabbit; goat; horse; cat; dog;
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;
KW ophthalmological; vulnerary; gene therapy; autoimmune disease; neoplasm;
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ss;
KW cerebrovascular disorder; nervous system disorder; bacterial infection;

KW fungal infection; viral infection; ocular disorder; endocrine disorder;
KW gastrointestinal disorder; renal disorder; respiratory disorder;
KW wound healing; skin aging; organ transplantation; food preservative;
KW tissue regeneration; anti-infertility; food additive.
XX
OS Homo sapiens.
XX
PN WO200155303-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01301.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.

PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-457723/49.
DR P-PSDB; AAU18577.
XX
PT Isolated polypeptide for treating, preventing and/or prognosing
PT respiratory disorders related to the lung including lung cancers and
PT also for testing and detection e.g. diagnosis -
XX
PS Claim 1; SEQ ID No 36; 507pp; English.
XX
CC Sequences AAS29839-AAS29930 represent cDNA molecules, which encode the
CC lung antigen polypeptides of the invention. Lung antigen polypeptides and
CC their associated polynucleotides are useful in the diagnosis, treatment
CC and prevention of various types of disorders in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological
CC condition can be determined by detecting the presence or absence of a
CC mutation in a lung antigen polynucleotide. The treatable disorders
CC include autoimmune diseases such as rheumatoid arthritis,
CC hyperproliferative disorders such as neoplasms of the breast or liver,
CC cardiovascular disorders such as cardiac arrest, cerebrovascular
CC disorders such as cerebral ischaemia, nervous system disorders such as
CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
CC ocular disorders such as corneal infection, endocrine disorders such as
CC premature labour and infertility, gastrointestinal disorders such as
CC Crohn's disease, renal disorders such as glomerulonephritis and
CC respiratory disorders such as asthma and pleurisy. The polypeptides can
CC also be used to aid wound healing, to prevent skin aging due to sunburn,
CC to maintain organs before transplantation, to regenerate tissues and in
CC chemotaxis. The polypeptides can also be used as a food additive or
CC preservative to increase or decrease storage capabilities.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO

Query Match 79.0%; Score 15.8; DB 22; Length 287;
Best Local Similarity 89.5%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120
```

RESULT 14
AAS28028
ID AAS28028 standard; cDNA; 287 BP.
XX
AC AAS28028;
XX
DT 07-NOV-2001 (first entry)
XX
DE Novel cDNA encoding for human respiratory antigen #160.
XX
KW Human; respiratory antigen; respiratory disorder; throat disorder;
KW lung disorder; nose disorder; lung cancer; gene therapy; cytostatic;
KW anti allergic; anti asthmatic; anti inflammatory; olfactory;
KW respiratory active; ss.
XX
OS Homo sapiens.
XX
PN WO200155448-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01333.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-476224/51.
DR P-PSDB; AAU17844.
XX
PT Isolated polypeptide for treating, preventing and/ or prognosing
PT disorders related to the respiratory system including respiratory
PT cancers and also for testing and detection e.g. diagnosis -
XX
PS Claim 4; SED ID No 170; 546pp; English.
XX
CC The present invention relates to the isolation of novel human
CC respiratory antigens (AAU17685-AAU17975), and cDNA and genomic
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful for preventing, treating and/or prognosing
CC disorders related to the respiratory system including throat
CC disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),
CC lung disorders e.g. pneumonia, allergic disorders e.g. asthma,
CC pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of
CC the respiratory tissues e.g. lung cancer. The polynucleotide sequences
CC of the invention are useful in gene therapy and antisense therapy.
CC AAS27869-AAS28159 encode for novel human respiratory antigens.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 287 BP; 85 A; 63 C; 28 G; 108 T; 3 other;

Query Match 79.0%; Score 15.8; DB 22; Length 287;
Best Local Similarity 89.5%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGCTAATTCGGTGCTACCT 19
||||| ||| |||||||||
Db 102 TGCTATTTCTGTGCTACCT 120

RESULT 15
ABZ52394/c
ID ABZ52394 standard; cDNA; 907 BP.
XX
AC ABZ52394;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide SEQ ID NO 1507.
XX
KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
KW expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN WO200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002WO-IB00890.
XX
PR 30-MAR-2001; 2001JP-0098371.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORQ) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
DR WPI; 2003-046817/04.
XX
PT Detection of expression of specific Aspergillus genes for monitoring
PT the fermentation and growth conditions of the fungus, using DNA probes
PT -
XX
PS Claim 1; SEQ ID NO 1507; 48pp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridising
CC to these sequences. The polynucleotides are useful for monitoring the

CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 907 BP; 249 A; 208 C; 205 G; 243 T; 2 other;

Query Match 79.0%; Score 15.8; DB 25; Length 907;
Best Local Similarity 89.5%; Pred. No. 91;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||  ||||||||||||||
Db        154 TGCTTCTTCGGTGCTACCT 136
```

Search completed: August 23, 2003, 10:38:19
Job time : 208.5 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 08:28:23 ; Search time 993 Seconds
(without alignments)
823.960 Million cell updates/sec

Title: US-09-927-585A-1
Perfect score: 20
Sequence: 1 tgtcaaaaaccgttgtgtcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AX301234	AX301234 Sequence
2	20	100.0	20	6	E05257	E05257 Primer for
3	20	100.0	20	6	E05247	E05247 Part of DNA
4	18.4	92.0	803	14	PPHE67	D10597 Human papil
5	18.4	92.0	7851	14	PPH35CG	M74117 Human papil
6	18.4	92.0	7879	14	HPV35H	X74477 Human papil
7	18.4	92.0	7912	14	PPH31A	J04353 Human papil
8	17.4	87.0	10909	1	AE013763	AE013763 Yersinia
c 9	17.4	87.0	102178	8	AP003803	AP003803 Oryza sat
c 10	17.4	87.0	223885	2	AC107604	AC107604 Rattus no
c 11	17.4	87.0	236021	2	AC094509	AC094509 Rattus no
c 12	17.4	87.0	236062	2	AC105688	AC105688 Rattus no
13	17.4	87.0	258050	1	AJ414153	AJ414153 Yersinia
14	16.8	84.0	20	6	E05251	E05251 Part of DNA
15	16.8	84.0	450	14	AF234529	AF234529 Human pap
16	16.8	84.0	450	14	AF234530	AF234530 Human pap
17	16.8	84.0	450	14	AF234531	AF234531 Human pap
18	16.8	84.0	450	14	AF478150	AF478150 Human pap
19	16.8	84.0	450	14	AF478151	AF478151 Human pap
20	16.8	84.0	450	14	AF478152	AF478152 Human pap
21	16.8	84.0	450	14	AF478153	AF478153 Human pap
22	16.8	84.0	450	14	AF478154	AF478154 Human pap
23	16.8	84.0	450	14	AF478155	AF478155 Human pap
24	16.8	84.0	450	14	AF478156	AF478156 Human pap
25	16.8	84.0	450	14	AF478157	AF478157 Human pap
26	16.8	84.0	450	14	AF478158	AF478158 Human pap
27	16.8	84.0	450	14	AF478159	AF478159 Human pap
28	16.8	84.0	450	14	AF478160	AF478160 Human pap
29	16.8	84.0	450	14	AF478161	AF478161 Human pap
30	16.8	84.0	450	14	AF478162	AF478162 Human pap
31	16.8	84.0	450	14	AF478163	AF478163 Human pap
32	16.8	84.0	450	14	AF478164	AF478164 Human pap
33	16.8	84.0	450	14	AF478165	AF478165 Human pap
34	16.8	84.0	450	14	AF478166	AF478166 Human pap
35	16.8	84.0	450	14	AF478167	AF478167 Human pap
c 36	16.8	84.0	766	6	AX186489	AX186489 Sequence
c 37	16.8	84.0	3443	6	AX536430	AX536430 Sequence
c 38	16.8	84.0	3840	8	YSCABOI	L31406 Saccharomyc
c 39	16.8	84.0	4569	8	SCYBL085W	Z35846 S.cerevisia
40	16.8	84.0	7801	14	D21208	D21208 Human papil
41	16.8	84.0	7812	14	AF131950	AF131950 Human pap
42	16.8	84.0	7824	14	PPH58	D90400 Human papil
43	16.8	84.0	7909	6	A07020	A07020 Human papil
44	16.8	84.0	7909	6	A12360	A12360 complete nu
45	16.8	84.0	7909	14	PPH33CG	M12732 Human papil

ALIGNMENTS

RESULT 1
AX301234
LOCUS AX301234 20 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 3 from Patent WO0185994.
ACCESSION AX301234
VERSION AX301234.1 GI:17382322
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Perseu,S., de Montis,A. and Floris,M.M.
TITLE Method and means for identifying hpv virus
JOURNAL Patent: WO 0185994-A 3 15-NOV-2001;
Bioanalisi Centro Sud S.N.C. Di Perseu Siniblado EC. (IT)

```
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="primer representative of region E6 of strains of
                     HPV virus having middle-high oncogenic hazard"
BASE COUNT        5 a      5 c      4 g      6 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 2
E05257
LOCUS       E05257                    20 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Primer for amplifying Human papillomavirus gene.
ACCESSION   E05257
VERSION     E05257.1  GI:2173447
KEYWORDS    JP 1993192200-A/27.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.
  TITLE     DETECTION OF HUMAN PAPILLOMA VIRUS
  JOURNAL   Patent: JP 1993192200-A 27 03-AUG-1993;
            TAKARA SHUZO CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1993192200-A/27
            PD   03-AUG-1993
            PF   19-AUG-1991 JP 1991230839
            PR   20-AUG-1990 JP 90P   217067
            PI   OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI
            FUKUSHIMA MICHIO,
            PI   FUJINAGA KEI
            PC   C12Q1/70,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        5 a      5 c      4 g      6 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 4.8;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 3
E05247
LOCUS       E05247                    20 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Part of DNA sequence of Human papillomavirus 31.
ACCESSION   E05247
VERSION     E05247.1  GI:2173437
KEYWORDS    JP 1993192200-A/17.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.
  TITLE     DETECTION OF HUMAN PAPILLOMA VIRUS
  JOURNAL   Patent: JP 1993192200-A 17 03-AUG-1993;
            TAKARA SHUZO CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            OS   Human papillomavirus
            PN   JP 1993192200-A/17
            PD   03-AUG-1993
            PF   19-AUG-1991 JP 1991230839
            PR   20-AUG-1990 JP 90P   217067
            PI   OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI
            FUKUSHIMA MICHIO,
            PI   FUJINAGA KEI
            PC   C12Q1/70,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        4 a      5 c      5 g      6 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 6;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 38;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||||||||||||
Db          1 TGTCAAAGACCGTTGTGTCC 20


RESULT 4
PPHE67
LOCUS       PPHE67                   803 bp    DNA     linear   VRL 02-FEB-1999
DEFINITION  Human papillomavirus gene for E6 and E7 proteins, complete cds.
ACCESSION   D10597 D01089
VERSION     D10597.1  GI:535788
KEYWORDS    E6 protein; E7 protein.
SOURCE      Human papillomavirus
  ORGANISM  Human papillomavirus
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   Fujinaga,K.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 803)
  AUTHORS   Fujinaga,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-FEB-1992) Kei Fujinaga, Cancer Research Institute,
            Sapporo Medical College, Department of Molecular Biology; S1, W17,
            Chuo-ku, Sapporo 060, Japan (Tel:011-611-2111, Fax:011-615-3099)
FEATURES             Location/Qualifiers
     source          1. .803
                     /organism="Human papillomavirus"
                     /mol_type="genomic DNA"
                     /isolate="Type Sapporo-A"
                     /db_xref="taxon:10566"
     CDS             52. .501
                     /codon_start=1
                     /product="E6 protein"
                     /protein_id="BAA01447.1"
                     /db_xref="GI:535789"
                     /translation="MFQDPAERPYKLHDLCNEVEESIHEICLNCVYCKQELQRSEVYD
                     FACYDLCIVYREGQPYGVCMKCLKFYSKISEYRRYRYSVYGETLEKQCNKQLCHLLIR
                     CITCQKPLCPVEKQRHLEEKKRFHNIGGRWTGRCMSCWKPTRRETEV"
```

```
     CDS             504. .803
                     /codon_start=1
                     /product="E7 protein"
                     /protein_id="BAA01448.1"
                     /db_xref="GI:535790"
                     /translation="MHGEITTLQDYVLDLEPEATDLYCYEQLCDSSEEEEDTIDGPAG
                     QAKPDTSNYNIVTSCCKCEATLRLCVQSTHIDIRKLEDLLMGTFGIVCPGCSQRA"
BASE COUNT      282 a    127 c    190 g    204 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 14;  Length 803;
  Best Local Similarity   95.0%;  Pred. No. 41;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||||| |||||||
Db        367 TGTCAAAAACCGCTGTGTCC 386


RESULT 5
PPH35CG
LOCUS       PPH35CG                 7851 bp    DNA     circular VRL 10-MAY-2002
DEFINITION  Human papillomavirus type 35 complete genome.
ACCESSION   M74117
VERSION     M74117.1  GI:333050
KEYWORDS    complete genome; major capsid protein; minor capsid protein;
            regulatory protein; replication protein; transformer protein.
SOURCE      Human papillomavirus type 35
  ORGANISM  Human papillomavirus type 35
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (sites)
  AUTHORS   Lorincz,A.T., Quinn,A.P., Lancaster,W.D. and Temple,G.F.
  TITLE     A New Type of Papillomavirus Associated with Cancer of the Uterine
            Cervix
  JOURNAL   Virology 159, 187-190 (1991)
  MEDLINE   87265470
REFERENCE   2  (bases 1 to 7851)
  AUTHORS   Marich,J.E., Pontsler,A.V., Rice,S.M., McGraw,K.A. and
            Dubensky,T.W.
  TITLE     The phylogenetic relationship and complete nucleotide sequence of
            human papillomavirus type 35
  JOURNAL   Virology 186 (2), 770-776 (1992)
  MEDLINE   92124753
   PUBMED   1310198
COMMENT     Original source text: Human papillomavirus type 35 cervical
            carcinoma DNA.
FEATURES             Location/Qualifiers
     source          1. .7851
                     /organism="Human papillomavirus type 35"
                     /proviral
                     /mol_type="genomic DNA"
                     /isolation_source="cervical carcinoma"
                     /db_xref="taxon:10587"
     5'UTR           join(7092. .7851,1. .109)
                     /standard_name="LCR"
                     /function="regulatory region"
     protein_bind    17. .22
                     /bound_moiety="Sp-1"
     CAAT_signal     complement(18. .26)
     protein_bind    24. .35
                     /note="putative"
                     /bound_moiety="E2"
                     /function="gene transcription"
     protein_bind    39. .50
                     /note="putative"
                     /bound_moiety="E2"
                     /function="gene transcription"
     TATA_signal     54. .59
                     /note="putative"
     gene            110. .559
                     /gene="E6"
     CDS             110. .559
                     /gene="E6"
                     /note="putative; E6 ORF from bp 59 to 559"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46966.1"
                     /db_xref="GI:333051"
                     /translation="MFQDPAERPYKLHDLCNEVEESIHEICLNCVYCKQELQRSEVYD
                     FACYDLCIVYREGQPYGVCMKCLKFYSKISEYRWYRYSVYGETLEKQCNKQLCHLLIR
                     CITCQKPLCPVEKQRHLEEKKRFHNIGGRWTGRCMSCWKPTRRETEV"
     misc_feature    230. .238
                     /gene="E6"
                     /standard_name="Splice donor"
     misc_feature    405. .416
                     /gene="E6"
                     /standard_name="Splice acceptor"
     gene            562. .861
                     /gene="E7"
     CDS             562. .861
                     /gene="E7"
                     /note="putative"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46967.1"
                     /db_xref="GI:333052"
                     /translation="MHGEITTLQDYVLDLEPEATDLYCYEQLCDSSEEEEDTIDGPAG
                     QAKPDTSNYNIVTSCCKCEATLRLCVQSTHIDIRKLEDLLMGTFGIVCPGCSQRA"
     gene            868. .2760
                     /gene="E1"
     CDS             868. .2760
                     /gene="E1"
                     /note="putative"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46968.1"
                     /db_xref="GI:333053"
                     /translation="MADPAGTDEGEGTGCNGWFFVEAVVSRRTGSSVEDENEDDCDRG
                     EDMVDFINDTDILNIQAETETAQALFHAQEEQTHKEAVQVLKRKYASSPLSSVSLCVN
                     NNISPRLKAICIENKNTAAKRRLFELPDSGYGNSEVEIHEIQQVEGHDTVEQCSMGSG
                     DSITSSSDERHDETPTRDIIQILKCSNANAAMLAKFKELFGISFTELIRPFKSDKSTC
                     TDWCVAAFGIAPSVANFKHITYVYIYNVYRVHGAMVILALLRFKVEKREQQLKTIDAK
                     LLCISAASMLIQPPKLRSTPAALYWFKTAMSNISEVDGETPEWIQRQTVLQHSFNDAI
                     FDLSEMVQWAYDNDFIDDSDIAYKYAQLAETNSNACAFLKSNSQAKIVKDCATMCRHY
                     KRAEKREMTMSQWIKRRCAQVDDDGDWRDIVRFLRYQQVDFVAFLSALKNFLHGVPKK
                     NCILIYGAPNTGKSLFGMSLMHFLQGAIISYVNSKSHFWLQPLYDAKIAMLDDATSPC
                     GIYRPIFKKCTRWKSYISFRCKALSIVHIMPTFTYYININAGKDDRWPYLHSRVVVFT
                     FHNEFPFDKNGNPEYGLNDKNWKSFFSRTWCRLNLHEEEVKENDGDAFPAFKCVSGQN
                     TRTLRD"
     gene            2693. .3796
                     /gene="E2"
     CDS             2693. .3796
                     /gene="E2"
                     /function="regulation of gene expression"
                     /note="putative"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46969.1"
                     /db_xref="GI:333054"
                     /translation="MMETLSQRLSVCQDKILEHYETDSTCLSDHIQYWKLIRLECAVF
                     YKAREMGIKTLNHQVVPTQAISKAKAMQAIELQLMLETLNTTEYSTEDWTLQETSIEL
                     YTTVPTRCLKKDVYTVEAQFDGDKQNTMHYTNWTHIYILEDSICTVVKGLVNYKGIYY
                     VHQGVETYYVTFREEAKKYGKKNIWEVHVGGQVIVCPESVFSSTELSTAEIATQLHAY
                     NTTETHTKACSVGTTETQKTNHKRLRGGTELPYNPTKRVRLSAVDSVDRGVYSTSDCT
                     NKDRCGSCSTTTPIVHLKGDANTLKCSRYRLGKYKALYQDASSTWRWTCTNDKKQIAI
                     VTLTYTTEYQRDKFLTTVKIPNTVTVSKGYMSI"
     gene            3273. .3563
                     /gene="E4"
     CDS             3273. .3563
                     /gene="E4"
                     /note="putative"
                     /codon_start=1
                     /product="unknown"
                     /protein_id="AAA46965.2"
                     /db_xref="GI:20522255"
```

```
                /translation="MFVLNLYLAAQNYPLLKLLHSYTPTTPPRPIPKPAPWAPQKPRR
                QITNDFEGVPSSPTTPPSECDSVPWTVLTEGSTLHLTAQTKTGVVVVVQLHL"
     gene       3793. .4038
                /gene="E5"
     CDS        3793. .4038
                /gene="E5"
                /note="putative"
                /codon_start=1
                /product="E5"
                /protein_id="AAA46970.1"
                /db_xref="GI:333056"
                /translation="MIDLTASSTVLLCFLLCFCVLLCLCLLVRSLLLSVSLYSALILL
                VLILWVTVATPLLAFVVSCFCIYLWMINAHAQYLAVQ"
     polyA_signal 4159. .4164
                /note="putative"
     gene       4184. .5593
                /gene="L2"
     CDS        4184. .5593
                /gene="L2"
                /note="putative; putative"
                /codon_start=1
                /product="minor capsid protein"
                /protein_id="AAA46971.1"
                /db_xref="GI:333057"
                /translation="MRHKRSTKRVKRASATQLYRTCKAAGTCPPDVIPKVEGNTVADQ
                ILKYGSMAVFFGGLGIGSGSGTGGRSGYVPLGTTPPTAATNIPIRPPVTVESIPLDTI
                GPLDSSIVSLVEETSFIESGAPVVTPRVPPTTGFTITTSTDTTPAILDVTSISTHDNP
                TFTDPSVLHPPTPAETSGHFVLSSSSISTHNYEEIPMDTFIVSTDSNNITNSTPIPGS
                RPTTRLGLYSKGTQQVKVVDPAFMTSPAKLITYDNPAYEGLNPDTTLQFEHEDISLAP
                DPDFMDIIALHRPALTSRKGTIRYSRVGNKRTMHTRSGKAIGARVHYYQDLSSITEDI
                ELQPLQHVPSSLPHTTVSTSLNDGMFDIYAPIDTEEDIIFSASSNNTLYTTSNTAYVP
                SNTTIPLSSGYDIPITAGPDIVFNSNTITNSVLPVPTGPIYSIIADGGDFYLHPSYYL
                LKRRRKAIPYFFADVSVAV"
     gene       5574. .7091
                /gene="L1"
     CDS        5574. .7091
                /gene="L1"
                /note="putative"
                /codon_start=1
                /product="major capsid protein"
                /protein_id="AAA46972.1"
                /db_xref="GI:333058"
                /translation="MSLWRSNEATVYLPPVSVSKVVSTDEYVTRTNIYYHAGSSRLLA
                VGHPYYAIKKQDSNKIAVPKVSGLQYRVFRVKLPDPNKFGFPDTSFYDPCLQRLVWAC
                TGVEVGRGQPLGVGISGHPLLNKLDDTENLNKYVGNSGNSGTDNRECISMDYKQTQLC
                LIGCRPPIGEHWGKGTPCNANQVKAGECPPLELLNTVLQDGDMVDTGFGAMDFTTLQA
                NKSDVPLDICSSICKYPDYLKMVSEPYGDMLFFYLRREQMFVRHLFNRAGTVGETVPA
                DLYIKGTTGTLPSTSYFPTPSGSMVTSDAQIFNKPYWLQRAQGHNNGICWSNQLFVTV
                VDTTRSTNMSVCSAVSSSSDSTYKNDNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYI
                HSMNPSILEDWNFGLTPPPSGTLEDTYRYVTSQAVTCQKPSAPKPKDDPLKNYTFWEV
                DLKEKFSADLDQFPLGRKFLLQAGLKARPNFRLGRRAAPASTSKKSSTKRRKVKS"
     repeat_region 7090. .7105
     repeat_region 7123. .7160
     protein_bind 7415. .7426
                /note="putative"
                /bound_moiety="E2"
                /function="gene transcription"
     protein_bind 7477. .7491
                /standard_name="glucocorticoid responsive element"
                /bound_moiety="hormone receptor"
     protein_bind 7514. .7519
                /bound_moiety="NF-1"
     protein_bind 7527. .7532
                /bound_moiety="NF-1"
     protein_bind 7534. .7539
                /bound_moiety="NF-1"
     enhancer   complement(7670. .7677)
                /standard_name="keratinocyte specific enhancer
                (CK-octomer)"
                /note="putative"
     protein_bind 7671. .7676
                /bound_moiety="NF-1"
     protein_bind 7695. .7700
                /bound_moiety="NF-1"
BASE COUNT     2553 a   1343 c   1568 g   2387 t

  Query Match             92.0%;  Score 18.4;  DB 14;  Length 7851;
  Best Local Similarity   95.0%;  Pred. No. 43;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||||| |||||||
Db        425 TGTCAAAAACCGCTGTGTCC 444


RESULT 6
HPV35H
LOCUS       HPV35H                  7879 bp    DNA     linear   VRL 24-FEB-1999
DEFINITION  Human papillomavirus type 35H genomic DNA.
ACCESSION   X74477
VERSION     X74477.1  GI:396997
KEYWORDS    E1 gene; E2 gene; E4 gene; E6 gene; E7 gene; early protein; L1
            gene; L2 gene; late protein.
SOURCE      Human papillomavirus type 35H
  ORGANISM  Human papillomavirus type 35H
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7879)
  AUTHORS   Delius,H. and Hofmann,B.
  TITLE     Primer-directed sequencing of human papillomavirus types
  JOURNAL   Curr. Top. Microbiol. Immunol. 186, 13-31 (1994)
  MEDLINE   94265501
   PUBMED   8205838
REFERENCE   2  (bases 1 to 7879)
  AUTHORS   Delius,H.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1993) H. Delius, Deutsches
            Krebsforschungszentrum, Abteilung ATV, Im Neuenheimer Feld 506, W
            6900 Heidelberg, FRG
FEATURES             Location/Qualifiers
     source          1. .7879
                     /organism="Human papillomavirus type 35H"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:31551"
                     /clone="insert in BamHI site of pBR322"
     gene            110. .559
                     /gene="E6"
     CDS             110. .559
                     /gene="E6"
                     /codon_start=1
                     /product="early protein"
                     /protein_id="CAA52561.1"
                     /db_xref="GI:396998"
                     /db_xref="SWISS-PROT:P27228"
                     /translation="MFQDPAERPYKLHDLCNEVEESIHEICLNCVYCKQELQRSEVYD
                     FACYDLCIVYREGQPYGVCMKCLKFYSKISEYRWYRYSVYGETLEKQCNKQLCHLLIR
                     CITCQKPLCPVEKQRHLEEKKRFHNIGGRWTGRCMSCWKPTRRETEV"
     gene            562. .861
                     /gene="E7"
     CDS             562. .861
                     /gene="E7"
                     /codon_start=1
                     /product="early protein"
                     /protein_id="CAA52562.1"
                     /db_xref="GI:396999"
                     /db_xref="SWISS-PROT:P27230"
                     /translation="MHGEITTLQDYVLDLEPEATDLYCYEQLCDSSEEEEDTIDGPAG
                     QAKPDTSNYNIVTSCCKCEATLRLCVQSTHIDIRKLEDLLMGTFGIVCPGCSQRA"
     gene            868. .2781
                     /gene="E1"
     CDS             868. .2781
                     /gene="E1"
                     /codon_start=1
                     /product="early protein"
                     /protein_id="CAA52563.1"
                     /db_xref="GI:397000"
```

```
                     /db_xref="SWISS-PROT:P27220"
                     /translation="MADPAGTDEGEGTGCNGWFFVEAVVSRRTGDPVSEDENEDDCDR
                     GEDMVDFINDTDILNIQAETETAQALFHAQEEQTHKEAVQVLKRKYASSPLSSVSLCV
                     NNNISPRLKAICIENKNTAAKRRLFELPDSGYGNSEVEIQQIQQVEGHDTVEQCSMGS
                     GDSITSSSDERHDETPTRDIIQILKCSNANAAMLAKFKELFGISFTELIRPFKSDKST
                     CTDWCVAAFGIAPSVAESLKTLIKPYCLYIHIQCLSCSWGMVILALLRFKCAKNRTTI
                     EKLLSKLLCISAASMLIQPPKLRSTPAALYWFKTAMSNISEVDGETPEWIQRQTVLQH
                     SFNDAIFDLSEMVQWAYDNDFIDDSDIAYKYAQLAETNSNACAFLKSNSQAKIVKDCA
                     TMCRHYKRAEKREMTMSQWIKRRCEKVDDDGDWRDIVRFLRYQQVDFVAFLSALKNFL
                     HGVPKKNCILIYGAPNTGKSLFGMSLMHFLQGAIISYVNSKSHFWLQPLYDAKIAMLD
                     DATSPCWAYIDQYLRNALDGNPISLDVKHKALVQLKCPPLLITSNINAGKDDRWPYLH
                     SRVVVFTFHNEFPFDKNGNPVYGLNDKNWKSFFSRTWCRLNLHEEEDKENDGDAFPAF
                     KCVSGQNTRTLRD"
     gene            2714. .2717
                     /gene="E2"
     CDS             2714. .>2717
                     /gene="E2"
                     /codon_start=1
                     /product="early protein"
                     /protein_id="CAA52564.1"
                     /db_xref="GI:584036"
                     /db_xref="SWISS-PROT:P27222"
                     /translation="M"
     gene            4211. .5620
                     /gene="L2"
     CDS             4211. .5620
                     /gene="L2"
                     /codon_start=1
                     /product="late protein"
                     /protein_id="CAA52565.1"
                     /db_xref="GI:397003"
                     /db_xref="SWISS-PROT:P27234"
                     /translation="MRHKRSTKRVKRASATQLYRTCKAAGTCPPDVIPKVEGNTVADQ
                     ILKYGSMAVFFGGLGIGSGSGTGGRSGYVPLGTTPPTAATNIPIRPPVTVESIPLDTI
                     GPLDSSIVSLVEETSFIESGAPVVTPRVPPTTGFTITTSTDTTPAILDVTSISTHDNP
                     TFTDPSVLHPPTPAETSGHFVLSSSSISTHNYEEIPMDTFIVSTDSNNITNSTPIPGS
                     RPTTRLGLYSKGTQQVKVVDPAFMTSPAKLITYDNPAYEGLNPDTTLQFEHEDISLAP
                     DPDFMDIIALHRPALTSRKGTIRYSRVGNKRTMHTRSGKAIGARVHYYQDLSSITEDI
                     ELQPLQHVPSSLPHTTVSTSLNDGMFDIYAPIDTEEDIIFSASSNNTLYTTSNTAYVP
                     SNTTIPLSSGYDIPITAGPDIVFNSNTITNTVLPVPTGPIYSIIADGGDFYLHPSYYL
                     LKRRRKRIPYFFADVSVAV"
     gene            5601. .7109
                     /gene="L1"
     CDS             5601. .7109
                     /gene="L1"
                     /codon_start=1
                     /product="late protein"
                     /protein_id="CAA52566.1"
                     /db_xref="GI:397004"
                     /db_xref="SWISS-PROT:P27232"
                     /translation="MSLWRSNEATVYLPPVSVSKVVSTDEYVTRTNIYYHAGSSRLLA
                     VGHPYYAIKKQDSNKIAVPKVSGLQYRVFRVKLPDPNKFGFPDTSFYDPASQRLVWAC
                     TGVEVGRGQPLGVGISGHPLLNKLDDTENSNKYVGNSGTDNRECISMDYKQTQLCLIG
                     CRPPIGEHWGKGTPCNANQVKAGECPPLELLNTVLQDGDMVDTGFGAMDFTTLQANKS
                     DVPLDICSSICKYPDYLKMVSEPYGDMLFFYLRREQMFVRHLFNRAGTVGETVPADLY
                     IKGTTGTLPSTSYFPTPSGSMVTSDAQIFNKPYWLQRAQGHNNGICWSNQLFVTVVDT
                     TRSTNMSVCSAVSSSDSTYKNDNFKEYLRHGEEYDLQFIFQLCKITLTADVMTYIHSM
                     NPSILEDWNFGLTPPPSGTLEDTYRYVTSQAVTCQKPSAPKPKDDPLKNYTFWEVDLK
                     EKFSADLDQFPLGRKFLLQAGLKARPNFRLGKRAAPASTSKKSSTKRRKVKS"
BASE COUNT     2570 a   1339 c   1570 g   2400 t
ORIGIN

  Query Match             92.0%;  Score 18.4;  DB 14;  Length 7879;
  Best Local Similarity   95.0%;  Pred. No. 43;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||| |||||||
Db        425 TGTCAAAAACCGCTGTGTCC 444


RESULT 7
PPH31A
```

```
LOCUS       PPH31A                  7912 bp    DNA     circular VRL 18-MAR-1994
DEFINITION  Human papillomavirus type 31 (HPV-31) complete genome.
ACCESSION   J04353
VERSION     J04353.1  GI:333048
KEYWORDS    complete genome.
SOURCE      Human papillomavirus type 31
  ORGANISM  Human papillomavirus type 31
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 7912)
  AUTHORS   Goldsborough,M.D., DiSilvestre,D., Temple,G.F. and Lorincz,A.T.
  TITLE     Nucleotide sequence of human papillomavirus type 31: a cervical
            neoplasia-associated virus
  JOURNAL   Virology 171 (1), 306-311 (1989)
  MEDLINE   89299478
   PUBMED   2545036
COMMENT     Original source text: Human papillomavirus type 31 DNA.
            Draft entry and computer-readable copy of sequence [1] kindly
            submitted by M.D.Goldsborough, 05-JUL-1989.
FEATURES             Location/Qualifiers
     source          1. .7912
                     /organism="Human papillomavirus type 31"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10585"
     TATA_signal     19. .24
     TATA_signal     69. .74
     gene            108. .557
                     /gene="E6"
     CDS             108. .557
                     /gene="E6"
                     /note="ORF E6 from bp 39 to 557"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46950.1"
                     /db_xref="GI:459916"
                     /translation="MFKNPAERPRKLHELSSALEIPYDELRLNCVYCKGQLTETEVLD
                     FAFTDLTIVYRDDTPHGVCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGICDLLIR
                     CITCQRPLCPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQV"
     misc_feature    228. .236
                     /gene="E6"
                     /standard_name="Splice donor"
     misc_feature    403. .414
                     /gene="E6"
                     /standard_name="Splice acceptor"
     gene            560. .856
                     /gene="E7"
     CDS             560. .856
                     /gene="E7"
                     /note="ORF E7 from bp 545 to 856"
                     /codon_start=1
                     /product="transforming protein"
                     /protein_id="AAA46951.1"
                     /db_xref="GI:459917"
                     /translation="MRGETPTLQDYVLDLQPEATDLHCYEQLPDSSDEEDVIDSPAGQ
                     AEPDTSNYNIVTFCCQCKSTLRLCVQSTQVDIRILQELLMGSFGIVCPNCSTRL"
     gene            862. .2751
                     /gene="E1"
     CDS             862. .2751
                     /gene="E1"
                     /note="ORF E1 from bp 850 to bp 2751"
                     /codon_start=1
                     /product="replication protein"
                     /protein_id="AAA46952.1"
                     /db_xref="GI:459918"
                     /translation="MADPAGTDGEGTGCNGWFYVEAVIDRQTGDNISEDENEDSSDTG
                     EDMVDFIDNCNVYNNQAEAETAQALFHAQEAEEHAEAVQVLKRKYVGSPLSDISSCVD
                     YNISPRLKAICIENNSKTAKRRLFELPDSGYGNTEVETQQMVQVEEQQTTLSCNGSDG
                     THSERENETPTRNILQVLKTSNGKAAMLGKFKELYGVSFMELIRPFQSNKSTCTDWCV
                     AAFGVTGTVAEGFKTLLQPYCLYCHLQSLACSWGMVMLMLVRFKCAKNRITIEKLLEK
                     LLCISTNCMLIQPPKLRSTAAALYWYRTGMSNISDVYGETPEWIERQTVLQHSFNDTT
                     FDLSQMVQWAYDNDVMDDSEIAYKYAQLADSDSNACAFLKSNSQAKIVKDCGTMCRHY
                     KRAEKRQMSMGQWIKSRCDKVSDEGDWRDIVKFLRYQQIEFVSFLSALKLFLKGVPKK
                     NCILIHGAPNTGKSYFGMSLISFLQGCIISYANSKSHFWLQPLADAKIGMLDDATTPC
```

```
                     WHYIDNYLRNALDGNPVSIDVKHKALMQLKCPPLLITSNINAGKDDRWPYLHSRLVVF
                     TFPNPFPFDKNGNPVYELSDKNWKSFFSRTWCRLNLHEEEDKENDGDSFSTFKCVSGQ
                     NIRTL"
     gene            2693. .3811
                     /gene="E2"
     CDS             2693. .3811
                     /gene="E2"
                     /note="ORF E2 from bp 2663 to 3811"
                     /codon_start=1
                     /product="regulatory protein"
                     /protein_id="AAA46953.1"
                     /db_xref="GI:459919"
                     /translation="METLSQRLNVCQDKILEHYENDSKRLCDHIDYWKHIRLECVLMY
                     KAREMGIHSINHQVVPALSVSKAKALQAIELQMMLETLNNTEYKNEDWTMQQTSLELY
                     LTAPTGCLKKHGYTVEVQFDGDVHNTMHYTNWKFIYLCIDGQCTVVEGQVNCKGIYYV
                     HEGHITYFVNFTEEAKKYGTGKKWEVHAGGQVIVFPPESVFSSDEISFAGIVTKLPTAN
                     NTTTSNSKTCALGTSEGVRRATTSTKRPRTEPEHRNTHHPNKLLRGDSVDSVNCGVIS
                     AAACTNQTRAVSCPATTPIIHLKGDANILKCLRYRLSKYKQLYEQVSSTWHWTCTDGK
                     HKNAIVTLTYISTSQRDDFLNTVKIPNTVSVSTGYMTI"
     gene            3270. .3578
                     /gene="E4"
     CDS             <3270. .3578
                     /gene="E4"
                     /note="ORF E4 from bp 3270 to 3578"
                     /codon_start=1
                     /protein_id="AAA46949.1"
                     /db_xref="GI:459915"
                     /translation="LFFLNLYLAVTKYPLLGLLQSYQQPTTPPHRIPKPAPWAPVKVC
                     GGRRRLLSDQEQSQSTETPTTPTSCCEATPWTVSTVGLSVQLHAQTKQGLSVVLQLHL
                     "
     gene            3816. .4070
                     /gene="E5"
     CDS             3816. .4070
                     /gene="E5"
                     /note="ORF E5 from bp 3804 to 4070"
                     /codon_start=1
                     /protein_id="AAA46954.1"
                     /db_xref="GI:459920"
                     /translation="MIELNISTVSIVLCFLLCFCVLLFVCLVIRPLVLSVSVYATLLL
                     LIVILWVIATSPLRCFCIYVVFIYIPLFVIHTHASFLSQQ"
     repeat_region   4099. .4134
     polyA_signal    4138. .4143
                     /note="putative"
     repeat_region   4143. .4158
     gene            4171. .5571
                     /gene="L2"
     CDS             4171. .5571
                     /gene="L2"
                     /note="ORF L2 from bp 4060 to bp 5571"
                     /codon_start=1
                     /product="minor capsid protein"
                     /protein_id="AAA46955.1"
                     /db_xref="GI:459921"
                     /translation="MRSKRSTKRTKRASATQLYQTCKAAGTCPSDVIPKIEHTTIADQ
                     ILRYGSMGVFFGGLGIGSGSGTGGRTGYVPLSTRPSTVSEASIPIRPPVSIDPVGPLD
                     PSIVSLVEESGIVDVGAPAPIPHPPTTSGFDIATTADTTPAILDVTSVSTHENPTFTD
                     PSVLQPPTPAETSGHLLLSSSSISTHNYEEIPMDTFIVSTNNENITSSTPIPGVRRPA
                     RLGLYSKATQQVKVIDPTFLSAPKQLITYENPAYETVNAEESLYFSNTSHNIAPDPDF
                     LDIIALHRPALTSRRNTVRYSRLGNKQTLRTRSGATIGARVHYYYDISSINPAGESIE
                     MQPLGASATTTSTLNDGLYDIYADTDFTVDTPATHNVSPSTAVQSTSAVSAYVPTNTT
                     VPLSTGFDIPIFSGPDVPIEHAPTQVFPFPLAPTTPQVSIFVDGGDFYLHPSYYMLKR
                     RRKRVSYFFTDVSVAA"
     gene            5552. .7066
                     /gene="L1"
     CDS             5552. .7066
                     /gene="L1"
                     /note="ORF L1 from bp 5516 to  7066"
                     /codon_start=1
                     /product="major capsid protein"
                     /protein_id="AAA46956.1"
                     /db_xref="GI:459922"
                     /translation="MSLWRPSEATVYLPPVPVSKVVSTDEYVTRTNIYYHAGSARLLT
                     VGHPYYSIPKSDNPKKIVVPKVSGLQYRVFRVRLPDPNKFGFPDTSFYNPETQRLVWA
                     CVGLEVGRGQPLGVGISGHPLLNKFDDTENSNRYAGGPGTDNRECISMDYKQTQLCLL
                     GCKPPIGEHWGKGSPCSNNAITPGDCPPLELKNSVIQDGDMVDTGFGAMDFTALQDTK
                     SNVPLDICNSICKYPDYLKMVAEPYGDTLFFYLRREQMFVRHFFNRSGTVGESVPTDL
                     YIKGSGSTATLANSTYFPTPSGSMVTSDAQIFNKPYWMQRAQGHNNGICWGNQLFVTV
                     VDTTRSTNMSVCAAIANSDTTFKSSNFKEYLRHGEEFDLQFIFQLCKITLSADIMTYI
                     HSMNPAILEDWNFGLTTPPSGSLEDTYRFVTSQAITCQKTAPQKPKEDPFKDYVFWEV
                     NLKEKFSADLDQFPLGRKFLLQAGYRARPKFKAGKRSAPSASTTTPAKRKKTKK"
     polyA_signal    7227. .7231
                     /note="putative"
     repeat_region   7291. .7302
     repeat_region   7314. .7333
                     /note="putative"
     protein_bind    7406. .7420
                     /standard_name="glucocorticoid responsive element"
                     /bound_moiety="hormone receptor"
     protein_bind    7477. .7488
                     /bound_moiety="E2"
                     /function="gene transcription"
     enhancer        7542. .7549
                     /standard_name="keratinocyte-dependant enhancer"
     protein_bind    7868. .7879
                     /bound_moiety="E2"
                     /function="gene transcription"
BASE COUNT     2528 a   1364 c   1572 g   2448 t
ORIGIN

  Query Match           92.0%;  Score 18.4;  DB 14;  Length 7912;
  Best Local Similarity  95.0%;  Pred. No. 43;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||||||||||||
Db        423 TGTCAAAGACCGTTGTGTCC 442


RESULT 8
AE013763
LOCUS       AE013763               10909 bp    DNA     linear   BCT 26-JUL-2002
DEFINITION  Yersinia pestis KIM section 163 of 415 of the complete genome.
ACCESSION   AE013763 AE009952
VERSION     AE013763.1  GI:21958388
KEYWORDS    .
SOURCE      Yersinia pestis KIM
  ORGANISM  Yersinia pestis KIM
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Yersinia.
REFERENCE   1  (bases 1 to 10909)
  AUTHORS   Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,
            Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,
            Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,
            Straley,S.C., McDonough,K.A., Nilles,M.L., Matson,J.S.,
            Blattner,F.R. and Perry,R.D.
  TITLE     Genome Sequence of Yersinia pestis KIM
  JOURNAL   J. Bacteriol. 184 (16), 4601-4611 (2002)
   PUBMED   12142430
REFERENCE   2  (bases 1 to 10909)
  AUTHORS   Deng,W., Burland,V., Plunkett,G. III, Boutin,A., Mayhew,G.F.,
            Liss,P., Perna,N.T., Rose,D.J., Mau,B., Zhou,S., Schwartz,D.C.,
            Fetherston,J.D., Lindler,L.E., Brubaker,R.R., Plana,G.V.,
            Straley,S.C., McDonough,K.A., Nilles,M.L., Matson,J.S.,
            Blattner,F.R. and Perry,R.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2002) Genetics, University of Wisconsin, 445
            Henry Mall, Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1. .10909
                     /organism="Yersinia pestis KIM"
                     /mol_type="genomic DNA"
                     /strain="KIM"
                     /db_xref="taxon:187410"
     gene            116. .1126
                     /gene="usg"
                     /note="synonym: y1598"
```

```
CDS             116. .1126
                /gene="usg"
                /function="putative"
                /note="contains semialdehyde dehydrogenase domains;
                residues 1 to 336 of 336 are 71.81 pct identical to
                residues 1 to 337 of 337 from E. coli K12 : B2319;
                residues 1 to 336 of 336 are 71.51 pct identical to
                residues 1 to 337 of 337 from GenPept : >gb|AAL21270.1|
                (AE008806) putative aspartate-semialdehyde dehydrogenase
                [Salmonella typhimurium LT2]"
                /codon_start=1
                /transl_table=11
                /product="putative enzyme"
                /protein_id="AAM85167.1"
                /db_xref="GI:21958389"
                /translation="MSDGWNIALLGATGAVGEALLELLNERQFPVGELYPLASERSAG
                ATVRFNGKSILVENVAEFDWSQAQLAFFVAGREASAQYAEEAGNAGCLVIDSSGLFAM
                EPDIPLVVPSVNPQVLADYRNRNIVAVADSLTSQLLTAIKPLTEQAGLSRLHVTNLIS
                ASAHGKAAVDDLAGQSAKLLNGIPAEPGIFPKQLAFNVLPLLADEEGSVSEERRLVDQ
                VRKVLQDEGLPITVSCIQSPVFYGHAQVVHLEALRPIAAEEARSELENCDDIQLSEED
                DYPTQVSDASGSDALSIGCVRNDYGIPEVLQFWSVADNIRFGGALMAIQTAERLLQEQ
                LY"
gene            1126. .2007
                /gene="truA"
                /note="synonym: y1599"
CDS             1126. .2007
                /gene="truA"
                /function="enzyme; aminoacyl tRNA synthetases, tRNA
                modification"
                /note="residues 20 to 290 of 293 are 78.22 pct identical
                to residues 4 to 270 of 270 from E. coli K12 : B2318;
                residues 20 to 290 of 293 are 78.59 pct identical to
                residues 4 to 270 of 270 from GenPept : >emb|CAD07600.1|
                (AL627274) tRNA pseudouridine synthase A [Salmonella
                enterica subsp. enterica serovar Typhi]"
                /codon_start=1
                /transl_table=11
                /product="pseudouridylate synthase I"
                /protein_id="AAM85168.1"
                /db_xref="GI:21958390"
                /translation="MSGFELDQEETLVQQNTLAQQEKPAEVAGLKIALGIEYNGSRYF
                GWQRQQEVASVQACLEAALSKVANEPIGVFCAGRTDAGVHATGQVVHFVTSAVRKDAA
                WTMGVNSHLPADIAVRWVKTVDNDFHARFSATARRYRYIIFSHRYRPAILAQGVTHCY
                MPLDAEKMERAAQCLLGENDFTSFRAVQCQSRTPWRNVKHVKVTRHGAYIVVDIKANA
                FVHHMVRNIVGSLIEIGCGNQDVTWMAELLALKDRTRAAATAKADGLYLVSVDYPDHF
                ALPKVPMGPLFLADDEG"
gene            2137. .2826
                /gene="dedA"
                /note="synonym: y1600"
CDS             2137. .2826
                /gene="dedA"
                /note="residues 10 to 225 of 229 are 85.64 pct identical
                to residues 1 to 216 of 219 from E. coli K12 : B2317;
                residues 10 to 225 of 229 are 85.18 pct identical to
                residues 1 to 216 of 219 from GenPept : >gb|AAL21268.1|
                (AE008806) putative DedA family [Salmonella typhimurium
                LT2]"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="AAM85169.1"
                /db_xref="GI:21958391"
                /translation="MCLFISRVFMEFIRFVIDFILNIDVHLAELVAQYGIWVYAILFL
                ILFCETGLVVTPFLPGDSLLFVAGALASLPSNDLNVHTMVTLMVIAAILGDAINYTIG
                RIFGEKLFSNPDSKIFRRSYLDKTHQFYEKHGGKAIILARFVPIIRTFAPFVAGMGKM
                SYRHFAAYNVIGALVWVLLFTYAGYLFGNVPIVQNNLKLLIVIIIVVSILPGVFEVWR
                HRRAAVRQKNQ"
gene            2901. .3971
                /gene="accD"
                /note="synonym: y1601"
CDS             2901. .3971
                /gene="accD"
                /function="enzyme; fatty acid and phosphatidic acid
                biosynthesis"
                /note="residues 53 to 350 of 356 are 86.24 pct identical
                to residues 1 to 296 of 304 from E. coli K12 : B2316;
                residues 53 to 350 of 356 are 86.91 pct identical to
                residues 1 to 296 of 304 from GenPept : >gb|AAL21267.1|
                (AE008806) acetylCoA carboxylase, beta subunit [Salmonella
                typhimurium LT2]"
                /codon_start=1
                /transl_table=11
                /product="acetyl CoA carboxylase, carboxytransferase
                component, beta subunit"
                /protein_id="AAM85170.1"
                /db_xref="GI:21958392"
                /translation="MWSVPADNPSFGALATSHWRVSQAFFIGEFSGRWLKHEQRTGNS
                QVQTERSSMSWIERILNKSNITQTRKASIPEGVWTKCDSCGQVLYRAELERNLEVCPK
                CDHHMRMSARARLHMLLDAGSEVELGSELEPKDILKFRDSKKYKDRISAAQKDTGEKD
                ALVAMKGTLQGMPIVAASFEFAFMGGSMASVVGARFVRAVEQALEDNCPLVCFSSSGG
                ARMQEALMSLMQMAKTSAALAKMQERGLPYISVLTDPTMGGVSASLAMLGDINIAEPK
                ALIGFAGPRVIEQTVREKLPPGFQRSEFLIEKGAIDMIVRRPVMRQTLASILSKLTHQ
                PQPSVVESKADTVAQPENQADV"
gene            4142. .5524
                /gene="folC"
                /note="synonym: y1602"
CDS             4142. .5524
                /gene="folC"
                /function="enzyme; biosynthesis of cofactors, carriers:
                Folic acid"
                /note="dihydrofolate synthetase; residues 31 to 458 of 460
                are 65.42 pct identical to residues 5 to 414 of 422 from
                E. coli K12 : B2315; residues 27 to 458 of 460 are 65.20
                pct identical to residues 1 to 414 of 422 from GenPept :
                >gb|AAL21266.1| (AE008806) multifunctional
                folylpolyglutamate synthase; dihydrofolate synthase, also
                has formylTHF polyglutamate synthase activity [Salmonella
                typhimurium LT2]"
                /codon_start=1
                /transl_table=11
                /product="dihydrofolate:folylpolyglutamate synthetase"
                /protein_id="AAM85171.1"
                /db_xref="GI:21958393"
                /translation="MGGTTGASVQATWSIAGIKPREVTGLMNNHQTPQATSPLAAWLC
                YLEHLHSQPIELGLERVKQVAERLDLLKPAPKIFTVAGTNGKGTTCCTLEAILLAAGL
                RVGVYSSPHLLRYTERVRIQGQELSEAEHSHSFAQIEAGRGDISLTYFEFGTLSALQL
                FKQAKLDVVILEVGLGGRLDATNIVDSDVAAITSIALDHTDWLGYDRESIGREKAGVF
                RGGKPAVVGEPDMPQSIADVAAELGAQLYRRDVAWKFSQQEPFDQQEPVDQQINGWHW
                QCGERQLTGLPVPNVPLANAATALAVLHYSELPLSDEAIRQGLQAASLPGRFQVVSEQ
                PLLILDVAHNPHAARYLVNRLAQVINPVNASKQGKVRAVVGMLSDKDIAGTLACLSER
                VDEWYCAPLEGPRGASAGQLAEHLVSARQFSDVETAWRQAMQDADTQDVVIVCGSFHT
                VAHVMAALHL"
gene            5638. .6366
                /gene="dedD"
                /note="synonym: y1603"
CDS             5638. .6366
                /gene="dedD"
                /function="putative membrane"
                /note="residues 12 to 239 of 242 are 55.45 pct identical
                to residues 1 to 209 of 211 from E. coli K12 : B2314;
                residues 3 to 239 of 242 are 56.72 pct identical to
                residues 1 to 218 of 220 from GenPept :
                >gb|AAG57443.1|AE005463_6 (AE005463) putative lipoprotein
                [Escherichia coli O157:H7 EDL933]"
                /codon_start=1
                /transl_table=11
                /product="putative lipoprotein"
                /protein_id="AAM85172.1"
                /db_xref="GI:21958394"
                /translation="MRVASQFQNRLVGTIILVALGVIVLPGLLDGKKKHYEDEFAAIP
                LVPKPGDSQEIDVVPPLNQPLPITPPEGAAQAVGAELNDDSVSQAANDAGNLAVEPTI
                VAPPPIQSKPVEVKPTVTKPVDVKPTETKPIETKPAEKKPLEVKPKPEIKPPVEEKAP
                VGQAYVVQLGALKNAAKVNEIVAMLRLSGHRAFTVPATPVQGEITRLYVGPDASKQKL
                QSALPELNALSGLNGQVKPYTTSR"
gene            6663. .7172
                /gene="cvpA"
```

```
                     /note="synonym: y1604"
     CDS             6663. .7172
                     /gene="cvpA"
                     /function="membrane; colicin-related functions"
                     /note="required for colicin V production; residues 1 to
                     160 of 169 are 83.12 pct identical to residues 1 to 160 of
                     162 from E. coli K12 : B2313; residues 1 to 160 of 169 are
                     83.12 pct identical to residues 1 to 160 of 162 from
                     GenPept : >gb|AAL21264.1| (AE008806) membrane protein
                     required for colicin V production [Salmonella typhimurium
                     LT2]"
                     /codon_start=1
                     /transl_table=11
                     /product="membrane protein"
                     /protein_id="AAM85173.1"
                     /db_xref="GI:21958395"
                     /translation="MVWIDYVIIGIIGFSALVSLIRGFVREALSLVTWGCAFFVASHF
                     YTYLAVYFTRFEDEIVRNGIAIGILFIATLIVGAIVNYVISSLVERTGLSGTDRVLGV

  Query Match             87.0%;  Score 17.4;  DB 1;  Length 10909;
  Best Local Similarity   94.7%;  Pred. No. 1.6e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTC 19
              ||||||||||||| ||||
Db       8159 TGTCAAAAACCGTTATGTC 8177


RESULT 9
AP003803/c
LOCUS       AP003803              102178 bp    DNA     linear   PLN 26-OCT-2002
DEFINITION  Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 7,
            BAC clone:OJ1060_D03.
ACCESSION   AP003803
VERSION     AP003803.3  GI:24414010
KEYWORDS    .
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
            clone:OJ1060_D03
  JOURNAL   Published Only in Database (2001)
REFERENCE   2  (bases 1 to 102178)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Takuji Sasaki, National Institute of
            Agrobiological Sciences, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     On Oct 26, 2002 this sequence version replaced gi:22775423.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), FGENESH
            (http://www.softberry.com/), GeneMark.hmm
            (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
            (http://www.tigr.org/tdb/glimmerm/glmr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), BLASTN and
            BLASTX. The genomic sequence was searched against NCBI NonRedundant
            Protein database, nr
            (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
            RGP. Protein homologies of the coding regions were searched against
            NCBI NonRedundant Protein database with BLASTP. ESTs represent the
            identified cDNA sequences using BLASTN with the corresponding DDBJ
            accession no. and RGP clone ID.
            A gene with identity or significant homology to a protein is
            classified based on the protein name to indicate the homology level
            such as same name, 'putative-' and '-like protein'.  A gene without
            significant homology to any protein but with EST homology (covering
            almost the entire length of partial sequence) is classified as an
            'unknown' protein. A gene predicted by two or more gene prediction
            programs is classified as a 'hypothetical' protein according to
            IRGSP standard. A gene predicted by a single gene prediction
            program is also classified as a probable 'hypothetical' protein and
            is included as a miscellaneous feature of the sequence.
            The orientation of the sequence is from -21M13 to M13rev of the BAC
            clone. This sequence of OJ1060_D03 clone has an overlap with
            OJ1127_E01 clone(DDBJ: AP003747)at 5' end and an overlap with
            OJ1113_E01(DDBJ: AP005437) at 3' end. The nucleotide sequence of
            this BAC clone was generated by combining Monsanto and RGP-Japan
            sequencing data. Detailed information on overlap and assembly
            quality together with annotation of this entry is available at
            http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES            Location/Qualifiers
     source         1. .102178
                    /organism="Oryza sativa (japonica cultivar-group)"
                    /mol_type="genomic DNA"
                    /cultivar="Nipponbare"
                    /db_xref="taxon:39947"
                    /chromosome="7"
                    /clone="OJ1060_D03"
     gene           1311. .1511
                    /gene="OJ1060_D03.1"
     misc_feature   1311. .1511
                    /gene="OJ1060_D03.1"
                    /note="This category is not included in IRGSP standard.
                    hypothetical ORF
                    predicted by GlimmerM"
     gene           complement(4270. .4578)
                    /gene="OJ1060_D03.2"
     CDS            complement(4270. .4578)
                    /gene="OJ1060_D03.2"
                    /note="hypothetical protein
                    predicted by GeneMark.hmm etc."
                    /codon_start=1
                    /protein_id="BAC22261.1"
                    /db_xref="GI:24414011"
                    /translation="MNQTFLRVRLGLPLTNCTIAGHISSLVHLQLSWHTPSTTFFPVL
                    HPSILATCDDGLGPFLDGASPPRLLLFLAAGADVVVDAGRAFSAGGRVYERSAASWRS
                    "
     gene           complement(join(5870. .6466,6570. .6647))
                    /gene="OJ1060_D03.3"
     CDS            complement(join(5870. .6466,6570. .6647))
                    /gene="OJ1060_D03.3"
                    /note="hypothetical protein
                    predicted by GeneMark.hmm etc."
                    /codon_start=1
                    /protein_id="BAC22262.1"
                    /db_xref="GI:24414012"
                    /translation="MERRLTPGVRAWSLHLRNMNDAAELSVRGRRLGGAAVKPCEEVY
                    VVGEDEMLHTISLTPPVIPLPSLSSSLSSLSLFFAGSAGTDGLVGSGAAAGDLGEAAP
                    SSPPSLSGDGLHAAPGPVASNDNDDMVGSCNCLWLKRHPLVCGSIIVINPVTGESLHI
                    PSPPLATSRSRHAGMISFGYHSTTGKYKIVHFPSNGGLVDEVTLGGTASSSSPPAMAA
                    ATVMGW"
     gene           join(8621. .8754,8874. .8983,9860. .9930,10812. .10932,
                    11030. .11337)
                    /gene="OJ1060_D03.4"
     CDS            join(8621. .8754,8874. .8983,9860. .9930,10812. .10932,
                    11030. .11337)
                    /gene="OJ1060_D03.4"
                    /note="mutator-like transposase-like protein"
                    /codon_start=1
                    /protein_id="BAC22263.1"
                    /db_xref="GI:24414013"
                    /translation="MRRRRRLHGAWRRCGASAACGSSSRPALQSGGSTFASLQNDDAV
                    GKIEDTYGRIGAYPNTYPYSFTYPIRIRSQPYQGILGARDACSICLRTTSPAAVESRH
                    IVAGASGSAHFPYMPTGDNVLNTPAWGLHITPRAPEQGHTQHTYDEYGSASTHHGVPA
                    YSPRTAFIEDFFSTDPPQEEVGMDYWHAAPQVTQPTQETEAGQGPDVTPQQTARDRHP
                    PDNLTYPTEQIRRRKKGGPSKRAKGTDRP"
     gene           12019. .12453
                    /gene="OJ1060_D03.5"
     CDS            12019. .12453
```

```
                /gene="OJ1060_D03.5"
                /note="hypothetical protein
                predicted by GlimmerM etc."
                /codon_start=1
                /protein_id="BAC22264.1"
                /db_xref="GI:24414014"
                /translation="MRCAMINHRLFDHARARVQSAHLLTVNTVPSPLLAALPNVRRCH
                LPNGEKKTPPKRSRCLHVSTSHRAKAATPMPITPYYYSSGGSNARGRGPSRSGARAAA
                AATWHGEEQTNGKKGDSLLPQPPKVKPPGAWCWWPAAHAMAP"
gene            join(12948. .14258,15290. .15376,16109. .16198)
                /gene="OJ1060_D03.6"
CDS             join(12948. .14258,15290. .15376,16109. .16198)
                /gene="OJ1060_D03.6"
                /note="arm repeat containing protein homolog-like"
                /codon_start=1
                /protein_id="BAC22265.1"
                /db_xref="GI:24414015"
                /translation="MVSLAGSQIPSPGQSPCAAARSQRRGAGYSMRTIRSALLQPDSC
                PGSPHVAAAYDAAGADSDMENLTDSVIDFHLSELAATAGPAHPAAVAKSSSANAAATE
                MLELSRDFSDYSSFNSDISGELERLAAAAAAVVTPRSDAPQVGAVDLNELESMDLSVE
                AAPLERVEPFVLACVRALGPDAAPDARRTAAARIRLLAKHRSDIRELIGVSGAIPALV
                PLLRSTDPVAQESAVTALLNLSLEERNRSAITAAGAIKPLVYALRTGTASAKQNAACA
                LLSLSGIEENRATIGACGAIPPLVALLSAGSTRGKKDALTTLYRLCSARRNKERAVSA
                GAVVPLIHLVGERGSGTSEKAMVVLASLAGIVEGRDAVVEAGGIPALVETIEDGPARE
                REFAVVALLQLCSECPRNRALLVREGAIPPLVALSQSGSARAKHKAETLLGYLREQRQ
                GGGGCRVEPVAASSLAFIVLRCVTFSLTPTSGSAAPSIWVTYSYT"
LTR             17638. .17869
                /note="5' LTR"
gene            complement(join(19591. .19658,20059. .20179))
                /gene="OJ1060_D03.7"
misc_feature    complement(join(19591. .19658,20059. .20179))
                /gene="OJ1060_D03.7"
                /note="This category is not included in IRGSP standard.
                hypothetical ORF
                predicted by GENSCAN"
gene            24046. .25584
                /gene="OJ1060_D03.8"
misc_feature    24046. .25584
                /gene="OJ1060_D03.8"
                /note="deletion
                pseudogene,Tam1 transposon protein TNP2"
                /pseudo
gene            complement(join(26731. .27006,27068. .27145))
                /gene="OJ1060_D03.9"
CDS             complement(join(26731. .27006,27068. .27145))
                /gene="OJ1060_D03.9"
                /note="hypothetical protein
                predicted by GENSCAN etc."
                /codon_start=1
                /protein_id="BAC22266.1"
                /db_xref="GI:24414016"
                /translation="MAARLRLGAAAVRRWCAAAASGRPGVRDVVASGGRIEVRVGIVG
                AAESCSAATTSAAASATRRPAAVWGEREIEVGERERGRRRRLSPQRCGGGGGSDDDAR
                EATARYDGGVGAGMP"
gene            28336. .31791
                /gene="OJ1060_D03.10"
misc_feature    28336. .31791
                /gene="OJ1060_D03.10"
                /note="probably inactive due to frameshift(s) in CDS
                probably inactive due to stop codon(s) in CDS
                pseudogene,hydroxyproline-rich glycoprotein"
                /pseudo
gene            36755. .37150
                /gene="OJ1060_D03.11"
CDS             36755. .37150
                /gene="OJ1060_D03.11"
                /note="hypothetical protein
                predicted by gm etc."
                /codon_start=1
                /protein_id="BAC22267.1"
                /db_xref="GI:24414017"
                /translation="MILFPPIHQTEAEHDGGRGKGQYPPVENSARACGRGGAPHATPT
                PRRHRAGRRGQPRPRSGLETRRFRLGKPKKYARLLIQRLQVDAPHWCGPSRAAATSSQ
                SPEASRWMPLLDFFGPYRSKPWKKNGSKI"
gene            38163. .39500
                /gene="OJ1060_D03.12"
CDS             38163. .39500
                /gene="OJ1060_D03.12"
                /note="contains ESTs
                AU183957(R0167),D46798(S11701),AU077610(E20630)
                similar to Arabidopsis thaliana chromosome5, At5g55530
                unknown protein"
                /codon_start=1
                /protein_id="BAC22268.1"
```

Query Match 87.0%; Score 17.4; DB 8; Length 102178;
Best Local Similarity 94.7%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 TGTCAAAAACCGTTGTGTC 19
              | ||||||||||||||||||
Db      84821 TTTCAAAAACCGTTGTGTC 84803
```

```
RESULT 10
AC107604/c
LOCUS       AC107604               223885 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-196K4, *** SEQUENCING IN PROGRESS
            ***.
ACCESSION   AC107604
VERSION     AC107604.4  GI:23266137
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 223885)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rivés,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
```

```
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 223885)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 223885)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 21, 2002 this sequence version replaced gi:21737824.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GPUL
                Center clone name: CH230-196K4
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 212075 bases at least Q40
                Consensus quality: 213656 bases at least Q30
                Consensus quality: 214464 bases at least Q20
                Estimated insert size: 222767; sum-of-contigs estimation
                Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   223885: contig of 223885 bp in length.
FEATURES            Location/Qualifiers
     source         1. .223885
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-196K4"
     misc_feature   1. .1816
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   1866. .2449
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH318237"
     misc_feature   complement(221124. .221882)
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BH318236"
     misc_feature   222550. .223885
                    /note="wgs_end_extension
                    clone_end:T7"
BASE COUNT    56904 a  48879 c  49011 g  60401 t   8690 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 2;  Length 223885;
  Best Local Similarity   94.7%;  Pred. No. 1.7e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              |||||||||| ||||||||
Db     102327 GTCAAAAACCTTTGTGTCC 102309


RESULT 11
AC094509/c
LOCUS       AC094509               236021 bp    DNA     linear   HTG 09-MAY-2003
DEFINITION  Rattus norvegicus clone CH230-4E17, *** SEQUENCING IN PROGRESS ***.
ACCESSION   AC094509
VERSION     AC094509.6  GI:30466778
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 236021)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
```

```
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 236021)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 236021)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 9, 2003 this sequence version replaced gi:23265564.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GAUD
                Center clone name: CH230-4E17
            -------------- Summary Statistics
                Assembly program: Atlas;
                Consensus quality: 218135 bases at least Q40
                Consensus quality: 221412 bases at least Q30
                Consensus quality: 223855 bases at least Q20
                Estimated insert size: 229824; sum-of-contigs estimation
                Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 1 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1   236021: contig of 236021 bp in length.
FEATURES            Location/Qualifiers
     source         1. .236021
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-4E17"
     misc_feature   1. .2200
                    /note="wgs_end_extension
                    clone_end:Sp6"
     misc_feature   5478. .6115
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH305262"
     misc_feature   complement(229315. .230090)
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BH305226"
     misc_feature   232901. .236021
                    /note="wgs_end_extension
                    clone_end:T7"
BASE COUNT    69961 a  46038 c  45434 g  64133 t  10455 others
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 2;  Length 236021;
  Best Local Similarity  94.7%;  Pred. No. 1.7e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              |||||||| ||||||||||
Db      74712 GTCAAAAAACGTTGTGTCC 74694


RESULT 12
AC105688/c
LOCUS       AC105688               236062 bp    DNA     linear   HTG 21-SEP-2002
DEFINITION  Rattus norvegicus clone CH230-11H4, *** SEQUENCING IN PROGRESS ***,
            2 unordered pieces.
ACCESSION   AC105688
VERSION     AC105688.3  GI:23264260
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 236062)
  AUTHORS   Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Loulseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
```

```
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
            Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 236062)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 236062)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-SEP-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Sep 21, 2002 this sequence version replaced gi:21743898.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequening reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/).  As a result, the
            sequence may extend beyond the ends of the clone and there may be
            contigs that consist entirely of whole genome shotgun sequence
            reads.  Both end sequences and whole genome shotgun sequence only
            contigs will be indicated in the feature table.
            -------------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
                Center project name: GNND
                Center clone name: CH230-11H4
            -------------- Summary Statistics
                Assembly program: Phrap; version 0.990329
                Consensus quality: 200479 bases at least Q40
                Consensus quality: 204791 bases at least Q30
                Consensus quality: 207411 bases at least Q20
                Estimated insert size: 251580; sum-of-contigs estimation
                Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
            --------------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
            * NOTE: This sequence may represent more than one clone.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 2 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1   234224: contig of 234224 bp in length
            *  234225   234324: gap of unknown length
            *  234325   236062: contig of 1738 bp in length.
FEATURES            Location/Qualifiers
     source         1. .236062
                    /organism="Rattus norvegicus"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:10116"
                    /clone="CH230-11H4"
     misc_feature   1064. .1913
                    /note="clone_boundary
                    clone_end:T7
                    site:EcoRI
                    end_sequence:BH341597"
     misc_feature   179267. .180672
                    /note="wgs_contig"
     misc_feature   complement(231465. .232270)
                    /note="clone_boundary
                    clone_end:Sp6
                    site:EcoRI
                    end_sequence:BH341599"
BASE COUNT    65632 a  42528 c  41770 g  59963 t  26169 others
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 2;  Length 236062;
  Best Local Similarity  94.7%;  Pred. No. 1.7e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              |||||||| ||||||||||
Db      63597 GTCAAAAAACGTTGTGTCC 63579


RESULT 13
AJ414153
LOCUS       AJ414153              258050 bp    DNA     linear   BCT 21-NOV-2002
DEFINITION  Yersinia pestis strain CO92 complete genome; segment 13/20.
ACCESSION   AJ414153 AL590842
VERSION     AJ414153.1  GI:15980576
KEYWORDS    .
SOURCE      Yersinia pestis CO92
  ORGANISM  Yersinia pestis CO92
            Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
            Enterobacteriaceae; Yersinia.
REFERENCE   1  (bases 1 to 258050)
  AUTHORS   Parkhill,J., Wren,B.W., Thomson,N.R., Titball,R.W., Holden,M.T.G.,
            Prentice,M.B., Sebaihia,M., James,K.D., Churcher,C., Mungall,K.L.,
            Baker,S., Basham,D., Bentley,S.D., Brooks,K., Cerdeno-Tarraga,A.M.,
            Chillingworth,T., Cronin,A., Davies,R.M., Davis,P., Dougan,G.,
            Feltwell,T., Hamlin,N., Holroyd,S., Jagels,K., Leather,S.,
            Karlyshev,A.V., Moule,S., Oyston,P.C.F., Quail,M., Rutherford,K.,
            Simmonds,M., Skelton,J., Stevens,K., Whitehead,S. and Barrell,B.G.
  TITLE     Genome sequence of Yersinia pestis, the causative agent of plague
  JOURNAL   Nature 413 (6855), 523-527 (2001)
  MEDLINE   21470413
   PUBMED   11586360
REFERENCE   2  (bases 1 to 258050)
  AUTHORS   Parkhill,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-OCT-2001) Submitted on behalf of the Yersinia
            sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
            Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT     Notes:
            Details of Y. pestis sequencing at the Sanger Centre are available
            on the World Wide Web.
            (URL, http://www.sanger.ac.uk/Projects/Y_pestis/).
FEATURES            Location/Qualifiers
     source         1. .258050
                    /organism="Yersinia pestis CO92"
                    /mol_type="genomic DNA"
                    /strain="CO92"
                    /db_xref="taxon:214092"
                    /note="biovar: Orientalis"
     gene           396. .590
                    /gene="YPO2590"
     CDS            396. .590
                    /gene="YPO2590"
                    /note="no significant database hits"
                    /codon_start=1
                    /transl_table=11
```

```
                /product="hypothetical protein"
                /protein_id="CAC92833.1"
                /db_xref="GI:15980577"
                /db_xref="SPTREMBL:Q8ZDH8"
                /translation="MSTLEKLVSAYCHTSLDFVASTVAFMESQKKNIDVDKIEAKLSV
                DERHFFRKRLTYYRDIYRPL"
     gene       complement(608. .790)
                /gene="YPO2591"
     CDS        complement(608. .790)
                /gene="YPO2591"
                /note="no significant database hits"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="CAC92834.1"
                /db_xref="GI:15980578"
                /db_xref="SPTREMBL:Q8ZDH7"
                /translation="MKTSHRAISTVREARKAIRSGYDKPIPKKGIKTGFKIDKPPDEM
                TTIEVHITQVAYLISV"
     gene       complement(972. .1937)
                /gene="YPO2592"
     CDS        complement(972. .1937)
                /gene="YPO2592"
                /note="Similar to Streptococcus pneumoniae transmembrane
                protein CAP33FM TR:O86896 (EMBL:AJ006986) (332 aa) fasta
                scores: E(): 9.7e-14, 25.4% id in 335 aa, and to internal
                region of Campylobacter jejuni probable enterochelin
                uptake permease CeuC TR:Q9PMU6 (EMBL:AL139078) (312 aa)
                fasta scores: E(): 0.13, 23.3% id in 296 aa"
                /codon_start=1
                /transl_table=11
                /product="putative membrane protein"
                /protein_id="CAC92835.1"
                /db_xref="GI:15980579"
                /db_xref="SPTREMBL:Q8ZDH6"
                /translation="MRNVGLHALKTLSCFSAVTFYSASKTCTEQCFLSGEVMGVLYFL
                SIIATPLFFMIIGYIDSNDKIDQKDIFRKLKSIITIIIFWNVLFYFINEDGFKRGYFL
                QSWLLFSIAIIYLINPIISKILKSNRTAIITLSCLVIFSVTIDLISAFSERPYLIDFP
                QYFRLWTWIFYYMTGRFLCSRICQKMTKLPKVRLAAKLLLLPTSISMYYYESFMSIYV
                YKTVNASYFLDNFHVLVLSLCLFVIFDNFDTQYEWIKKTLAYISPSMIGVYILHDGIF
                YFIASAYNLSDVTLRFTLLLSVFAASVLLSRALLLNKYTSRFISF"
     misc_feature complement(1005. .1070)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1101. .1166)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1206. .1262)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1308. .1361)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1401. .1451)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1497. .1547)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1569. .1634)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1665. .1715)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     misc_feature complement(1755. .1820)
                /gene="YPO2592"
                /note="one of 9 probable transmembrane helices predicted
                for YPO2592 by TMHMM2.0"
     gene       2360. .2977
                /gene="YPO2593"
     CDS        2360. .2977
                /gene="YPO2593"
                /note="Similar to Salmonella typhimurium fimbriae W
                protein FimW SW:FIMW_SALTY (P37928) (198 aa) fasta scores:
                E(): 0.0066, 21.9% id in 178 aa, and to C-terminus of
                Erwinia amylovora activator of exopolysaccharide
                synthesis, RcsB TR:P96320 (EMBL:Y09848) (215 aa) fasta
                scores: E(): 0.0003, 43.7% id in 71 aa"
                /codon_start=1
                /transl_table=11
                /product="putative LuxR-family regulatory proteins"
                /protein_id="CAC92836.1"
                /db_xref="GI:15980580"
                /db_xref="SPTREMBL:Q8ZDH5"
                /translation="MGIAKMISCMIVDNDKYQTLGMVSIVKKIFTSLGFKKEIKFYRK
                AFYSADVIFIGVDEFSFFEALKRLDKTPAEADVFLICDARLNSFLQGIPRFSDVTMIF
                REDSLETVSNKIATVFKRKFRGFKDNLIERKTAKVLHLVSSERQYLTPNENIVLKLFN
                EGFSGGDIARILKKSEKTVSGQKRSAMKKLGARTDVELIKMFMFK"
     misc_feature 2792. .2968
                /gene="YPO2593"
                /note="Pfam match to entry PF00196 GerE, Bacterial
                regulatory proteins, luxR family, score 21.90, E-value
                0.00026"
     misc_feature 2846. .2911
                /gene="YPO2593"
                /note="Predicted helix-turn-helix motif with score 1182
                (+3.21 SD) at aa 163-184, sequence FSGGDIARILKKSEKTVSGQKR"
     gene       complement(3297. .3668)
                /gene="YPO2594"
     CDS        complement(3297. .3668)
                /gene="YPO2594"
                /note="no significant database hits"
                /codon_start=1
                /transl_table=11
                /product="hypothetical protein"
                /protein_id="CAC92837.1"
                /db_xref="GI:15980581"
                /db_xref="SPTREMBL:Q8ZDH4"
                /translation="MIKKGAEVTFEPTTNTKGLSAFAVKVAIESKYIFIANERIKLTS
                IKSFNTFTKEVPAQAEVDKANTILSVNLLMNKIRPQEEDISEKTIPLKMLSITTFQNV
                TYTFSEHEVDIDSTVAKLKSI"
     gene       4054. .4263
                /gene="cspE"
                /note="synonym: YPO2595"
     CDS        4054. .4263
                /gene="cspE"
                /note="Similar to Escherichia coli probable cold shock
                protein CspE SW:CSPE_ECOLI (P36997) (68 aa) fasta scores:
                E(): 1e-25, 94.1% id in 68 aa, and to Vibrio cholerae
                hypothetical protein TR:AAC72388 (EMBL:AF098260) (69 aa)
                fasta scores: E(): 3.1e-26, 94.2% id in 69 aa. Similar to
                YPO1746 (84.1% identity in 69 aa overlap), YPO1655 (83.3%
                identity in 66 aa overlap), YPO3644 (73.6% identity in 72
                aa overlap), YPO3643 (72.2% identity in 72 aa overlap),
                YPO2659 (63.9% identity in 72 aa overlap) and YPO1398
                (53.6% identity in 69 aa overlap)."
                /codon_start=1
                /transl_table=11
                /product="putative cold shock protein"
                /protein_id="CAC92838.1"
                /db_xref="GI:15980582"
                /db_xref="SPTREMBL:Q8ZDH3"
                /translation="MSKIKGSVKWFNESKGFGFITPEDGSKDVFVHFSAIASNGFKTL
                AEGQRVEFEITNGAKGPSAANVIAI"
     misc_feature 4060. .4260
                /gene="cspE"
                /note="Pfam match to entry PF00313 CSD, 'Cold-shock'
```

```
                    DNA-binding domain, score 152.30, E-value 8.6e-42"
     misc_feature   4102. .4161
                    /gene="cspE"
                    /note="PS00352 'Cold-shock' DNA-binding domain signature."
     gene           complement(4686. .5069)
                    /gene="crcB"
                    /note="synonym: YPO2596"
     CDS            complement(4686. .5069)
                    /gene="crcB"

  Query Match             87.0%;  Score 17.4;  DB 1;  Length 258050;
  Best Local Similarity   94.7%;  Pred. No. 1.7e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTC 19
              ||||||||||||||| ||||
Db     194210 TGTCAAAAACCGTTATGTC 194228


RESULT 14
E05251
LOCUS       E05251                    20 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Part of DNA sequence of Human papillomavirus 33.
ACCESSION   E05251
VERSION     E05251.1  GI:2173441
KEYWORDS    JP 1993192200-A/21.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.
  TITLE     DETECTION OF HUMAN PAPILLOMA VIRUS
  JOURNAL   Patent: JP 1993192200-A 21 03-AUG-1993;
            TAKARA SHUZO CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            OS   Human papillomavirus
            PN   JP 1993192200-A/21
            PD   03-AUG-1993
            PF   19-AUG-1991 JP 1991230839
            PR   20-AUG-1990 JP 90P   217067
            PI   OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI
            FUKUSHIMA MICHIO,
            PI   FUJINAGA KEI
            PC   C12Q1/70,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No.
FEATURES            Location/Qualifiers
     source         1. .20
                    /organism="synthetic construct"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:32630"
BASE COUNT       4 a      5 c      4 g      7 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 6;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 3.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||| ||||||||
Db          1 TGTCAAAGACCTTTGTGTCC 20


RESULT 15
AF234529
LOCUS       AF234529                 450 bp    DNA     linear   VRL 28-MAR-2000
DEFINITION  Human papillomavirus type 58 strain E6HK1 E6 protein gene, complete
            cds.
ACCESSION   AF234529
VERSION     AF234529.1  GI:7331167
KEYWORDS    .
SOURCE      Human papillomavirus type 58
  ORGANISM  Human papillomavirus type 58
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 450)
  AUTHORS   Chan,P.K.S., Lam,C.W., Li,W.H., Chan,M.Y.M., Cheung,J.L.K. and
            Cheng,A.F.
  TITLE     Nucleotide sequence variation of the E6 gene of human
            papillomavirus type 58 from Chinese women with normal cervices and
            cancerous lesions
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 450)
  AUTHORS   Chan,P.K.S., Lam,C.W., Li,W.H., Chan,M.Y.M., Cheung,J.L.K. and
            Cheng,A.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-FEB-2000) Department of Microbiology, The Chinese
            University of Hong Kong, Prince of Wales Hospital, Shatin, NT.,
            Hong Kong
FEATURES            Location/Qualifiers
     source         1. .450
                    /organism="Human papillomavirus type 58"
                    /mol_type="genomic DNA"
                    /strain="E6HK1"
                    /isolate="QEH61"
                    /db_xref="taxon:10598"
                    /country="China"
                    /note="from cervical scrape of patient with invasive
                    squamous cell carcinoma of the cervix"
     CDS            1. .450
                    /note="oncoprotein; transforming; trans-activating;
                    p53-binding"
                    /codon_start=1
                    /product="E6 protein"
                    /protein_id="AAF60305.1"
                    /db_xref="GI:7331168"
                    /translation="MFQDAEEKPRTLHDLCQALETSVHEIELKCVECKKTLQRSEVYD
                    FVFADLRIVYRDGNPFAVCKVCLRLLSKISEYRHYNYSLYGDTLEQTLKKCLNEILIR
                    CIICQRPLCPQEKKRHVDLNKRFHNISGRWTGRCAVCWRPRRRQTQV"
BASE COUNT      158 a     60 c    104 g    128 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 14;  Length 450;
  Best Local Similarity   90.0%;  Pred. No. 3.2e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||| ||||||||
Db        316 TGTCAAAGACCATTGTGTCC 335


Search completed: August 23, 2003, 11:11:33
Job time : 995 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 05:06:07 ; Search time 206.5 Seconds
(without alignments)
261.447 Million cell updates/sec

Title: US-09-927-585A-1
Perfect score: 20
Sequence: 1 tgtcaaaaaccgttgtgtcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
25: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	14	AAQ48566	HPV E6/7 region pr
	2	20	100.0	20	24	AAS18612	Middle-high hazard
	3	20	100.0	20	25	ABZ75144	Oncogenic HPV regi
	4	18.4	92.0	20	14	AAQ48556	HPV E6/7 region pr
	5	18.4	92.0	1023	14	AAQ48575	HPV E6/7 Genomic r
	6	16.8	84.0	20	14	AAQ48560	HPV E6/7 region pr
c	7	16.8	84.0	540	24	ABT11435	Yeast selected int
c	8	16.8	84.0	541	24	ABS62861	Selected Interacti
c	9	16.8	84.0	766	22	AAH70910	Human cervical can
	10	16.8	84.0	774	10	AAN91602	Partial nucleotide
c	11	16.8	84.0	3443	24	ABQ76303	S. cerevisiae BAX-
c	12	16	80.0	530	22	AAF25493	Nucleotide sequenc
c	13	16	80.0	1447	22	AAF25498	Nucleotide sequenc
c	14	16	80.0	265118	22	AAH41227	Pyrococcus abyssi
c	15	15.8	79.0	611	24	ABQ55587	Human ovarian anti
c	16	15.8	79.0	1409	25	ABT19257	Aspergillus fumiga
c	17	15.8	79.0	1410	25	ABT21077	Aspergillus fumiga
c	18	15.8	79.0	1518	25	ABT20479	Aspergillus fumiga
c	19	15.8	79.0	1521	25	ABT18663	Aspergillus fumiga
c	20	15.8	79.0	3518	25	ABT19883	Aspergillus fumiga
c	21	15.8	79.0	3520	25	ABT18069	Drosophila melanog
c	22	15.8	79.0	29046	23	ABL10478	Pyrococcus abyssi
c	23	15.8	79.0	349980	22	AAH41225	DNA encoding novel
c	24	15.4	77.0	3339	23	AAS68508	Human immune/haema
c	25	15.4	77.0	13309	22	AAK83980	HPV E6/7 region pr
	26	15.2	76.0	20	14	AAQ48549	hPV Primer set 13,
c	27	15.2	76.0	29	16	AAQ76065	E6 oncogene from H
c	28	15.2	76.0	100	19	AAV39991	HPV-PTM related ol
c	29	15.2	76.0	138	24	ABQ73533	HPV-PTM related ol
c	30	15.2	76.0	193	24	ABQ73531	Human papillomavir
	31	15.2	76.0	456	19	ABQ76223	Human foetal liver
	32	15.2	76.0	477	22	ABA58702	Probe #6139 for ge
	33	15.2	76.0	477	22	ABA27673	Human brain expres
	34	15.2	76.0	477	22	AAK06838	Human bone marrow
	35	15.2	76.0	477	22	AAK32556	Probe #6055 for ge
	36	15.2	76.0	477	22	AAI16122	Probe #7067 used t
	37	15.2	76.0	477	22	AAI38381	Human liver single
	38	15.2	76.0	477	23	ABS32267	Human genome-deriv
	39	15.2	76.0	477	24	ABS07346	E6 siRNA sequence.
	40	15.2	76.0	477	25	AAL54430	HPV16 E6 siRNA der
	41	15.2	76.0	477	25	AAL54445	Human papilloma vi
	42	15.2	76.0	519	17	AAT31834	Mammalian vestibul
	43	15.2	76.0	525	22	AAS05471	HPV16 E6/E7 encodi
	44	15.2	76.0	570	16	AAQ75470	Drosophila melanog
c	45	15.2	76.0	642	23	ABL04975	

ALIGNMENTS

RESULT 1
AAQ48566
ID AAQ48566 standard; DNA; 20 BP.
XX
AC AAQ48566;
XX
DT 22-FEB-1994 (first entry)
XX
DE HPV E6/7 region probe.
XX
KW Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.
XX
OS Synthetic.
XX
PN JP05192200-A.
XX
PD 03-AUG-1993.
XX
PF 19-AUG-1991; 91JP-0230839.
XX
PR 20-AUG-1990; 90JP-0217067.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
DR WPI; 1993-277497/35.
XX
PT Detecting benign and/or malignant human papilloma virus - by
PT detecting DNA sequence of E6 and/or E7 region of human papilloma
PT virus
XX
PS Disclosure; Page 15; 18pp; Japanese.

XX
CC The probe is used to detect benign and/or malignant human papilloma
CC virus. The probe binds to the E6 and/or E7 region of the virus.
XX
SQ Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||||||||||||||||||||
Db 1 TGTCAAAAACCGTTGTGTCC 20

RESULT 2
AAS18612
ID AAS18612 standard; DNA; 20 BP.
XX
AC AAS18612;
XX
DT 26-FEB-2002 (first entry)
XX
DE Middle-high hazard human papillomavirus (HPV) E6, PCR primer.
XX
KW Human papillomavirus; HPV; early gene; E6; PCR primer;
KW middle-high hazard; ss.
XX
OS Human papillomavirus.
XX
PN WO200185994-A2.
XX
PD 15-NOV-2001.
XX
PF 02-MAY-2001; 2001WO-IB00771.
XX
PR 05-MAY-2000; 2000IT-MO00091.
XX
PA (BIOA-) BIOANALISI CENT SUD DI PERSEU SINIBLADO.
XX
PI Perseu S, De Montis A, Floris MM;
XX
DR WPI; 2002-075253/10.
XX
PT Identifying viral DNA of human papilloma virus in cellular material
PT collected from an individual, by using polymerase and identifying
PT sequences complementary to regions E of the viral genome using a
PT reagent -
XX
PS Claim 9; Page 12; 17pp; English.
XX
CC The invention describes a novel method of identifying viral DNA of human
CC papillomavirus (HPV) in cellular material collected from an individual.
CC The method comprises extracting DNA from the cellular material,
CC amplifying the DNA by using a polymerase, and identifying sequences
CC complementary to sequences in the regions E (early genes) e.g. E6 and E7
CC of the HPV genome using an identifying reagent. The presence of HPV in
CC cells is now based on identification of a nucleotide sequence in the L1
CC region of the genome. Identification of E6, E7 and L1 make it possible to
CC identify the presence of HPV virus in every stage of infection and to
CC distinguish between virus strains having low hazard and virus strains
CC having middle/high hazard. This is the primer used to identify the E6
CC region of middle-high hazard HPV strains in infected cells, described in
CC the method of the invention.
XX
SQ Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

Query Match 100.0%; Score 20; DB 24; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||||||||||||||||||||
Db 1 TGTCAAAAACCGTTGTGTCC 20

RESULT 3
ABZ75144
ID ABZ75144 standard; DNA; 20 BP.
XX
AC ABZ75144;
XX
DT 30-MAY-2003 (first entry)
XX
DE Oncogenic HPV region E6-E7 PCR primer VI-3, SEQ ID NO:18.
XX
KW Reaction mixture; dessicated; dried; enzymatic; stabilised;
KW condensation reaction inhibitor; polymer mesh; nucleic acid manipulation;
KW amplification; sequencing; hybridisation; restriction analysis;
KW diagnosis; ready-to-use; hot start reaction; HPV; identification;
KW oncogenic type; E6-E7; PCR; primer; ss.
XX
OS Human papillomavirus type 16.
OS Human papillomavirus type 18.
OS Human papillomavirus type 31.
OS Human papillomavirus type 33.
OS Human papillomavirus type 35.
OS Human papillomavirus type 52b.
OS Human papillomavirus type 58.
XX
PN WO200272002-A2.
XX
PD 19-SEP-2002.
XX
PF 11-MAR-2002; 2002WO-ES00109.
XX
PR 12-MAR-2001; 2001ES-0000569.
XX
PA (BIOT-) BIOTOOLS BIOTECHNOLOGICAL & MEDICAL LAB.
XX
PI Franco De Sarabia Rosado PM, Limones Lopez G, Madejon Seiz A;
PI Marin Alberdi MD;
XX
DR WPI; 2003-067358/06.
XX
PT Stabilized reaction mixture containing an enzyme, useful for performing
PT nucleic acid reactions, includes three-component stabilizing mixture
PT and is at least partially dried -
XX
PS Example VI; Page 41; 76pp; Spanish.
XX
CC The invention relates to a method for preparing a stabilised reaction
CC mixture which contains at least one enzyme and which is at least
CC partially dried. The method involves removing the water from an aqueous
CC reaction mixture containing an enzyme and a three-component stabiliser
CC solution until the residual moisture content is 30% or less. The
CC three-component stabiliser solution used in the method comprises an
CC agent which protects against dessication, an inhibitor of condensation
CC reactions between carbonyl or carboxy groups and amino or phosphate
CC groups, and an inert polymer that forms a mesh structure that inhibits
CC the mobility of the dried reactants. The dried reaction mixtures can be
CC used as "ready-to-use" mixtures for performing a wide range of nucleic
CC acid manipulations such as amplification, sequencing, hybridisation
CC and/or restriction analysis, e.g., for the diagnostic detection of
CC pathogens or mutations. The dried reaction mixtures can contain all the
CC components needed to perform a particular reaction, already deposited in
CC a reaction vessel, which eliminates the need for multiple additions
CC (which are sources of errors and contamination), thereby improving
CC repeatability and reliability. The dried reaction mixtures can be
CC transported and stored at ambient temperature without significant loss
CC of activity, and are suitable for "hot start" reactions. Sequences
CC ABZ75142-ABZ75145 represent PCR primers used in an exemplification to
CC amplify regions of human papillomavirus (HPV) genomes in order to
CC characterise their type. Primers ABZ75142-ABZ75143 are PCR primers used
CC to amplify a 450 bp region designated L1 which is shared by all HPV
CC types, while PCR primers ABZ75144-ABZ75145 are used to amplify a 250 bp

```
CC   region designated E6-E7 from oncogenic HPV types only.
XX
SQ   Sequence 20 BP; 5 A; 5 C; 4 G; 6 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 25;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 4
AAQ48556
ID   AAQ48556 standard; DNA; 20 BP.
XX
AC   AAQ48556;
XX
DT   22-FEB-1994  (first entry)
XX
DE   HPV E6/7 region probe.
XX
KW   Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.
XX
OS   Synthetic.
XX
PN   JP05192200-A.
XX
PD   03-AUG-1993.
XX
PF   19-AUG-1991;   91JP-0230839.
XX
PR   20-AUG-1990;   90JP-0217067.
XX
PA   (TAKI ) TAKARA SHUZO CO LTD.
XX
DR   WPI; 1993-277497/35.
XX
PT   Detecting benign and/or malignant human papilloma virus - by
PT   detecting DNA sequence of E6 and/or E7 region of human papilloma
PT   virus
XX
PS   Disclosure; Page 14; 18pp; Japanese.
XX
CC   The probe is used to detect benign and/or malignant human papilloma
CC   virus. The probe binds to the E6 and/or E7 region of the virus.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 5 G; 6 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 14;  Length 20;
  Best Local Similarity   95.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||||||||||||
Db          1 TGTCAAAGACCGTTGTGTCC 20


RESULT 5
AAQ48575
ID   AAQ48575 standard; DNA; 1023 BP.
XX
AC   AAQ48575;
XX
DT   22-FEB-1994  (first entry)
XX
DE   HPV E6/7 Genomic region.
XX
KW   Human papilloma virus; HPV; benign; malignant; E6; E7; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             250..696
FT                   /*tag=  a
FT                   /label= ORF-1
FT                   /note= "E6 region"
FT   CDS             702..998
FT                   /*tag=  b
FT                   /label= ORF-2
FT                   /note= "E7 region"
XX
PN   JP05192200-A.
XX
PD   03-AUG-1993.
XX
PF   19-AUG-1991;   91JP-0230839.
XX
PR   20-AUG-1990;   90JP-0217067.
XX
PA   (TAKI ) TAKARA SHUZO CO LTD.
XX
DR   WPI; 1993-277497/35.
DR   P-PSDB; AAR40919, AAR48202.
XX
PT   Detecting benign and/or malignant human papilloma virus - by
PT   detecting the DNA sequence of E6 and/or E7 region of human papilloma
PT   virus
XX
PS   Claim 1; Page 16-17; 18pp; Japanese.
XX
CC   Probes were designed to detect benign and/or malignant human papilloma
CC   virus. The probes bind to the E6 and/or E7 region of human papilloma
CC   virus.
XX
SQ   Sequence 1023 BP; 355 A; 170 C; 227 G; 271 T; 0 other;

  Query Match             92.0%;  Score 18.4;  DB 14;  Length 1023;
  Best Local Similarity   95.0%;  Pred. No. 18;
  Matches   19;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||||| |||||||
Db        565 TGTCAAAAACCGCTGTGTCC 584


RESULT 6
AAQ48560
ID   AAQ48560 standard; DNA; 20 BP.
XX
AC   AAQ48560;
XX
DT   22-FEB-1994  (first entry)
XX
DE   HPV E6/7 region probe.
XX
KW   Human papilloma virus; HPV; E6; E7; benign; malignant; probe; ss.
XX
OS   Synthetic.
XX
PN   JP05192200-A.
XX
PD   03-AUG-1993.
XX
PF   19-AUG-1991;   91JP-0230839.
XX
PR   20-AUG-1990;   90JP-0217067.
XX
PA   (TAKI ) TAKARA SHUZO CO LTD.
XX
DR   WPI; 1993-277497/35.
XX
PT   Detecting benign and/or malignant human papilloma virus - by
PT   detecting DNA sequence of E6 and/or E7 region of human papilloma
PT   virus
```

```
XX
PS   Disclosure; Page 15; 18pp; Japanese.
XX
CC   The probe is used to detect benign and/or malignant human papilloma
CC   virus. The probe binds to the E6 and/or E7 region of the virus.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 4 G; 7 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 14;  Length 20;
  Best Local Similarity   90.0%;  Pred. No. 86;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| ||| ||||||||
Db          1 TGTCAAAGACCTTTGTGTCC 20


RESULT 7
ABT11435/c
ID   ABT11435 standard; DNA; 540 BP.
XX
AC   ABT11435;
XX
DT   10-DEC-2002  (first entry)
XX
DE   Yeast selected interacting domain coding sequence SEQ ID NO: 369.
XX
KW   Yeast; protein-protein interaction; Selected Interacting Domain;
KW   SID (RTM); secretion yield; cancer; neurodegenerative disease; fungicide;
KW   cytostatic; neuroprotective; gene; ds.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200266504-A2.
XX
PD   29-AUG-2002.
XX
PF   14-FEB-2002; 2002WO-EP02299.
XX
PR   16-FEB-2001; 2001US-269266P.
XX
PA   (HYBR-) HYBRIGENICS.
XX
PI   Legrain P;
XX
DR   WPI; 2002-674913/72.
DR   P-PSDB; ABJ11118.
XX
PT   New protein-protein complexes of Saccharomyces cerevisiae, useful in
PT   drug screening or development, for developing yeast strains with better
PT   secretion yield of protein, or in gene therapy (e.g. to treat Candida
PT   infection or cancer)
XX
PS   Claim 7; Page 168; 357pp; English.
XX
CC   The present invention relates to complexes between Saccharomyces
CC   cerevisiae Selected Interacting Domain (SID (RTM)) proteins and coding
CC   sequences. The protein complexes of S. cerevisiae are useful in drug
CC   development, in screening drugs or agents that modulate the interaction
CC   of proteins, for developing yeast strains with better secretion yield of
CC   protein, and in gene therapy. The protein complexes, polypeptides and
CC   polynucleotides are useful for preventing or treating Candida infection,
CC   cancer or neurodegenerative diseases in humans or animals. The present
CC   sequence is a coding sequence of the invention.
XX
SQ   Sequence 540 BP; 184 A; 116 C; 113 G; 127 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 24;  Length 540;
  Best Local Similarity   90.0%;  Pred. No. 1.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||| ||||||||||||| |||
Db        309 TGTGAAAAACCGTTGTTTCC 290


RESULT 8
ABS62861/c
ID   ABS62861 standard; DNA; 541 BP.
XX
AC   ABS62861;
XX
DT   05-NOV-2002  (first entry)
XX
DE   Selected Interacting Domain (SID) polynucleotide #58.
XX
KW   Yeast; selected interacting domain; SID; gene; ds; antifungal; cancer;
KW   cytostatic; neuroprotective; Candida infection; gene therapy;
KW   neurodegenerative disease.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200259255-A2.
XX
PD   01-AUG-2002.
XX
PF   25-JAN-2002; 2002WO-EP01350.
XX
PR   26-JAN-2001; 2001US-264577P.
XX
PA   (HYBR-) HYBRIGENICS.
XX
PI   Legrain P;
XX
DR   WPI; 2002-619165/66.
DR   P-PSDB; ABG77247.
XX
PT   New complex between two interacting bait and prey Saccharomyces
PT   cerevisiae polypeptides, useful for preventing or treating Candida
PT   infection, cancer or neurodegenerative diseases in a mammal
XX
PS   Claim 7; Page 80; 196pp; English.
XX
CC   The invention relates to a complex between two interacting Saccharomyces
CC   cerevisiae polypeptides, comprising two Selected Interacting Domain (SID)
CC   polypeptides as bait and prey proteins. A pharmaceutical composition
CC   comprising the complex is useful for preventing or treating Candida
CC   infection, cancer and neurodegenerative diseases in a human or animal,
CC   preferably in a mammal. This sequence represents DNA encoding a SID
CC   polypeptide of the invention.
XX
SQ   Sequence 541 BP; 184 A; 116 C; 114 G; 127 T; 0 other;

  Query Match             84.0%;  Score 16.8;  DB 24;  Length 541;
  Best Local Similarity   90.0%;  Pred. No. 1.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||| ||||||||||||| |||
Db        310 TGTGAAAAACCGTTGTTTCC 291


RESULT 9
AAH70910/c
ID   AAH70910 standard; cDNA; 766 BP.
XX
AC   AAH70910;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2184.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
```

PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
PS Claim 1; Page 462; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 766 BP; 197 A; 158 C; 159 G; 244 T; 8 other;

Query Match 84.0%; Score 16.8; DB 22; Length 766;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||||||| || |||||||
Db 734 TGTCAAAAGCCCTTGTGTCC 715

RESULT 10
AAN91602
ID AAN91602 standard; DNA; 774 BP.
XX
AC AAN91602;
XX
DT 17-JUL-1990 (first entry)
XX
DE Partial nucleotide sequence (5' end) of human papilloma virus (HPV)
DE type 33 (HPV-33).
XX
KW Human papilloma virus; type 33; in situ hybridisation assay;
KW cellular smear; cervical carcinoma.
XX
OS Human papilloma virus.
XX
PN WO8902934-A.
XX
PD 06-APR-1989.
XX
PF 30-SEP-1988; 88WO-US03367.
XX
PR 02-OCT-1987; 87US-0103979.
XX
PA (MICR-) MICROPROBE CORP.
XX
PI Schwartz DE, Adams TH;
XX
DR WPI; 1989-114406/15.
XX
PT Hybridisation test for human papilloma virus in cell smears -
PT by reaction with long labelled probe specific for particular
PT virus types, esp. for examining cervical smears
XX
PS Disclosure; ; 39pp; English.
XX
CC The patent is for a rapid in situ hybridisation assay for detecting and
CC typing human papilloma virus (HPV) in non-frozen cellular smears fixed to
CC a support in absence of aldehyde-based crosslinking reagents. The assay
CC comprises: (1) combining nucleic acid in the sample with at least one
CC detectable probe able to hybridise with 1 or more HPV types; and (2)
CC detecting presence or absence of hybrid complexes. Opt. several probes
CC are used, eg one for HPV types 6 and 11, associated with benign warts,
CC and one for types 16, 18, 31, 33 and 35, associated with cervical
CC cancer. The assay can differentiate between HPV types. It is esp. used
CC as a secondary test. The probes can be synthesised or cloned.
XX
SQ Sequence 774 BP; 270 A; 132 C; 163 G; 209 T; 0 other;

Query Match 84.0%; Score 16.8; DB 10; Length 774;
Best Local Similarity 90.0%; Pred. No. 1.1e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||||||| ||| |||||||
Db 424 TGTCAAAGACCTTTGTGTCC 443

RESULT 11
ABQ76303/c
ID ABQ76303 standard; cDNA; 3443 BP.
XX
AC ABQ76303;
XX
DT 21-NOV-2002 (first entry)
XX
DE S. cerevisiae BAX-associated cDNA fragment SEQ ID 31.
XX
KW Bax; Bax-resistance; cytostatic; fungicide; immunosuppressive; virucide;
KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
KW neurodegeneration; cell death; ss.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200264766-A2.
XX
PD 22-AUG-2002.
XX
PF 21-DEC-2001; 2001WO-EP15398.
XX
PR 22-DEC-2000; 2000EP-0870318.
PR 04-JAN-2001; 2001EP-0870002.
PR 09-JAN-2001; 2001EP-0870003.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
XX
DR WPI; 2002-667002/71.
DR P-PSDB; ABG93037.
XX
PT New isolated nucleic acid representing a synthetic Bax gene, useful as
PT medicament for treating, preventing and/or alleviating yeast or fungal
PT infections or proliferative disorders, or for preventing apoptosis in
PT certain diseases -
XX
PS Claim 36; Figure 1; 344pp; English.
XX
CC This invention describes a novel nucleic acid representing a synthetic
CC Bax gene. The Bax gene of the invention is useful for identifying
CC Bax-resistant yeast or fungi, identifying, or obtaining and identifying

CC Candida spp. sequences that are differentially expressed in a pathway
CC eventually leading to programmed cell death or identifying inhibitors or
CC inhibitor sequences of Bax-induced cell death. The products of the
CC invention have cytostatic, fungicide; immunosuppressive, virucide and
CC vasotropic activity and can be used in vaccines or for gene therapy. The
CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
CC antisense molecules and antibodies are useful as medicaments or in
CC preparing a medicament for treating, preventing and/or alleviating
CC diseases associated with yeast or fungi or proliferative disorders, such
CC as cancer, or for preventing apoptosis in certain diseases. The compounds
CC or polypeptides , or the genetically modified organism are useful for
CC preparing a medicament for modifying the endogenic flora of humans and
CC other mammals. The vaccine is useful for immunising against yeast or
CC fungal infections. Apoptosis-related diseases include autoimmune disease,
CC ischaemia, diseases related with viral infections or neurodegenerations.
CC This sequence represents a polynucleotide associated with the Bax gene
CC described in the disclosure of the invention.
XX
SQ Sequence 3443 BP; 1229 A; 718 C; 650 G; 846 T; 0 other;

Query Match 84.0%; Score 16.8; DB 24; Length 3443;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||| |||||||||||| |||
Db 2900 TGTGAAAAACCGTTGTTTCC 2881

RESULT 12
AAF25493/c
ID AAF25493 standard; DNA; 530 BP.
XX
AC AAF25493;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a rice AMP deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT CDS 202..312
FT /*tag= a
FT /product= "AMP deaminase"
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
DR WPI; 2001-159866/16.
DR P-PSDB; AAB31953.
XX
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
PS Claim 2; Page 53; 72pp; English.
XX
CC The present sequence encodes an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX
SQ Sequence 530 BP; 156 A; 123 C; 98 G; 150 T; 3 other;

Query Match 80.0%; Score 16; DB 22; Length 530;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGT 16
||||||||||||||||
Db 50 TGTCAAAAACCGTTGT 35

RESULT 13
AAF25498/c
ID AAF25498 standard; DNA; 1447 BP.
XX
AC AAF25498;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a soybean adenosine deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
KW inosine; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 188..1255
FT /*tag= a
FT /product= "adenosine deaminase"
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PF 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
DR WPI; 2001-159866/16.
DR P-PSDB; AAB31958.
XX
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
PS Claim 2; Page 67-68; 72pp; English.
XX
CC The present sequence encodes an andenosine deaminase. The specification
CC also describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and

```
CC   catabolism of adenosine and AMP. In humans, this may lead to death of
CC   white blood cells, which causes severe immunodeficiencies. The AMP
CC   deaminase and adenosine deaminase may be used to prepare antibodies
CC   to these proteins, and to design or identify herbicides that inhibit
CC   their enzyme activities. The polynucleotides are used as probes for
CC   genetically and physically mapping genes that they compose, and as
CC   markers for traits linked to those genes, where such information may be
CC   used in plant breeding to develop lines with desired phenotypes. The
CC   nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC   homologous proteins from the same or other plant species, and in
CC   polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC   acid fragments encoding homologous genes from DNA or RNA. These may
CC   also be used to create transgenic plants in which the polypeptides are
CC   overexpressed or suppressed, and as probes in direct fluorescent in
CC   situ hybridisation (FISH).
XX
SQ   Sequence 1447 BP; 442 A; 283 C; 324 G; 398 T; 0 other;

  Query Match             80.0%;  Score 16;  DB 22;  Length 1447;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGT 16
              ||||||||||||||||
Db         57 TGTCAAAAACCGTTGT 42


RESULT 14
AAH41227/c
ID   AAH41227 standard; DNA; 265118 BP.
XX
AC   AAH41227;
XX
DT   29-OCT-2001  (first entry)
XX
DE   Pyrococcus abyssi genomic fragment #6.
XX
KW   Hyperthermophilic archaeon; hyperthermophilic protein; ds.
XX
OS   Pyrococcus abyssi.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..49980
FT                   /*tag= a
FT                   /note= "This sequence overlaps with the 3' end of
FT                   AAH41226"
XX
PN   FR2792651-A1.
XX
PD   27-OCT-2000.
XX
PF   21-APR-1999;   99FR-0005034.
XX
PR   21-APR-1999;   99FR-0005034.
XX
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PA   (IFRE-) IFREMER INST FR RECH EXPL MER.
XX
PI   Forterre P,  Thierry JC,  Prieur D,  Dietrich J,  Lecompte O;
PI   Querellou J,  Weissenbach J,  Saurin W,  Heilig R;
XX
DR   WPI; 2001-126236/14.
XX
PT   New nucleotide sequences isolated from Pyrococcus abyssi encode
PT   proteins useful in industry -
XX
PS   Claim 1; Page 593-665; 1657pp; French.
XX
CC   The present invention relates to the genomic sequence of Pyrococcus
CC   abyssi and P. abyssi proteins (see AAB96053-AAB96842). P. abyssi is a
CC   hyperthermophilic archaeon, which is isolated from deep-sea hydrothermal
CC   vents. The present sequence is a fragment of the genomic sequence of P.
CC   abyssi. The 5' end of this sequence overlaps with the 3' end of AAH41226.
CC   The proteins of the present invention have various potential industrial
CC   uses, since the proteins are stable at very high temperatures, some up to
CC   110 degrees centigrade.
CC   Note: This patent is in the same patent family as WO200065062, which
CC   contains additional sequences as shown in AAB99132-AAB99143,
CC   AAH75903-AAH75920 and AAG66436.
XX
SQ   Sequence 265118 BP; 75226 A; 61302 C; 54005 G; 74585 T; 0 other;

  Query Match             80.0%;  Score 16;  DB 22;  Length 265118;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGT 16
              ||||||||||||||||
Db     231084 TGTCAAAAACCGTTGT 231069


RESULT 15
ABQ55587/c
ID   ABQ55587 standard; cDNA; 611 BP.
XX
AC   ABQ55587;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HODFV22 cDNA, SEQ ID NO:1467.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   P-PSDB; ABP42510.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 1; SEQ ID No 1467; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
```

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents cDNA encoding a human ovarian antigen of the
CC invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 611 BP; 212 A; 106 C; 120 G; 154 T; 19 other;

Query Match 79.0%; Score 15.8; DB 24; Length 611;
Best Local Similarity 89.5%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 GTCAAAAACCGTTGTGTCC 20
              |||||| ||| ||||||||
Db        291 GTCAAATACCTTTGTGTCC 273
```

Search completed: August 23, 2003, 10:38:17
Job time : 208.5 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:18:48 ; Search time 50 Seconds
(without alignments)
176.553 Million cell updates/sec

Title: US-09-927-585A-2
Perfect score: 20
Sequence: 1 tgctaattcggtgctacctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	20	100.0	30	1	US-08-479-817-9	Sequence 9, Appli
c	2	20	100.0	30	1	US-08-461-038-9	Sequence 9, Appli
c	3	20	100.0	30	1	US-08-461-645-9	Sequence 9, Appli
	4	20	100.0	1107	4	US-09-000-094-19	Sequence 19, Appl
	5	20	100.0	1128	4	US-09-000-094-21	Sequence 21, Appl
	6	20	100.0	1398	4	US-09-000-094-23	Sequence 23, Appl
	7	20	100.0	4770	4	US-09-000-094-45	Sequence 45, Appl
	8	20	100.0	8010	3	US-09-521-526-2	Sequence 2, Appli
	9	20	100.0	8010	5	PCT-US95-11859-2	Sequence 2, Appli
c	10	15.4	77.0	3340	4	US-09-228-986-5	Sequence 5, Appli
	11	14.8	74.0	692	4	US-09-345-236B-76	Sequence 76, Appl
	12	14.8	74.0	692	4	US-09-345-236B-78	Sequence 78, Appl
c	13	14.8	74.0	692	4	US-09-345-236B-80	Sequence 80, Appl
	14	14.8	74.0	1285	4	US-09-149-476-80	Sequence 80, Appl
	15	14.8	74.0	11970	3	US-09-345-217-1	Sequence 1, Appli
	16	14.8	74.0	152331	3	US-09-128-155-16	Sequence 16, Appl
	17	14.4	72.0	747	3	US-09-393-877-8	Sequence 8, Appli
	18	14.4	72.0	747	4	US-09-703-807-8	Sequence 8, Appli
c	19	14.2	71.0	747	4	US-08-487-283A-17	Sequence 17, Appl
c	20	14.2	71.0	747	4	US-08-487-283A-21	Sequence 21, Appl
c	21	14.2	71.0	747	4	US-08-487-283A-22	Sequence 22, Appl
c	22	14.2	71.0	747	4	US-08-487-283A-23	Sequence 23, Appl
c	23	14.2	71.0	750	4	US-08-487-283A-11	Sequence 11, Appl
c	24	14.2	71.0	750	4	US-08-487-283A-16	Sequence 16, Appl
c	25	14.2	71.0	820	3	US-08-990-823-54	Sequence 54, Appl
c	26	14.2	71.0	820	4	US-09-477-135A-54	Sequence 54, Appl
c	27	14.2	71.0	2616	4	US-09-134-001C-1142	Sequence 1142, Ap
	28	14.2	71.0	3943	1	US-08-369-796-3	Sequence 3, Appli
	29	14.2	71.0	3943	2	US-08-852-091-3	Sequence 3, Appli
	30	14.2	71.0	3943	2	US-08-820-754-3	Sequence 3, Appli
	31	14.2	71.0	3943	3	US-08-956-652-3	Sequence 3, Appli
	32	14.2	71.0	3943	3	US-08-956-869-3	Sequence 3, Appli
	33	14.2	71.0	3943	3	US-08-948-547-3	Sequence 3, Appli
	34	14.2	71.0	3943	3	US-09-364-970-10	Sequence 10, Appl
	35	14.2	71.0	3943	4	US-08-956-653A-3	Sequence 3, Appli
	36	14.2	71.0	3943	5	PCT-US95-17025-3	Sequence 3, Appli
	37	14.2	71.0	4003	3	US-09-087-465-1	Sequence 1, Appli
	38	14.2	71.0	4003	4	US-09-972-800A-1	Sequence 1, Appli
	39	14.2	71.0	4608	3	US-09-041-886-24	Sequence 24, Appl
	40	14.2	71.0	4608	5	PCT-US94-05277-1	Sequence 1, Appli
	41	14.2	71.0	7739	3	US-09-195-966-1	Sequence 1, Appli
	42	14.2	71.0	7745	1	US-08-299-675-1	Sequence 1, Appli
	43	14.2	71.0	7745	1	US-08-485-241-1	Sequence 1, Appli
	44	14.2	71.0	7745	2	US-08-874-162-1	Sequence 1, Appli
c	45	14.2	71.0	34094	4	US-09-292-034-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-479-817-9/c
; Sequence 9, Application US/08479817
; Patent No. 5597910
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,817
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/307,026
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/805,537
;      FILING DATE:  11-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3440
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-817-9

  Query Match           100.0%;  Score 20;  DB 1;  Length 30;
  Best Local Similarity  100.0%;  Pred. No. 0.051;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 2
US-08-461-038-9/c
; Sequence 9, Application US/08461038
; Patent No. 5610017
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,038
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3451
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-461-038-9

  Query Match             100.0%;  Score 20;  DB 1;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 3
US-08-461-645-9/c
; Sequence 9, Application US/08461645
; Patent No. 5686244
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,645
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-461-645-9

  Query Match             100.0%;  Score 20;  DB 1;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 0.051;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 4
US-09-000-094-19
; Sequence 19, Application US/09000094
; Patent No. 6365160
;   GENERAL INFORMATION:
;        APPLICANT: WEBB, Elizabeth Ann
;                   MARGETTS, Mary Brigid
;                   COX, John Cooper
;                   FRAZER, Ian
;                   MCMILLAN, Nigel Alan John
;                   WILLIAMS, Mark Philip
;                   MOLONEY, Margaret Bridget
;                   Holland
;                   EDWARDS, Stirling John
;        TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
;        NUMBER OF SEQUENCES: 50
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: FOLEY & LARDNER
;             STREET: 3000 K Street, N.W.
;             CITY: Washington
;             STATE: D.C.
;             COUNTRY: U.S.A.
;             ZIP: 20007-5109
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/000,094
;             FILING DATE: 21-Apr-1998
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: WO PCT/AU96/00473
;             FILING DATE: 26-JUL-1996
;             APPLICATION NUMBER: AU PN 4439/95
;             FILING DATE: 27-JUL-1995
;        ATTORNEY/AGENT INFORMATION:
```

```
;            NAME: BENT, Stephen A.
;            REGISTRATION NUMBER: 29,768
;            REFERENCE/DOCKET NUMBER: 017227/0137
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 672-5300
;            TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 19:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1107 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..1104
;       SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-000-094-19

  Query Match             100.0%;  Score 20;  DB 4;  Length 1107;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 5
US-09-000-094-21
; Sequence 21, Application US/09000094
; Patent No. 6365160
;   GENERAL INFORMATION:
;       APPLICANT: WEBB, Elizabeth Ann
;                  MARGETTS, Mary Brigid
;                  COX, John Cooper
;                  FRAZER, Ian
;                  MCMILLAN, Nigel Alan John
;                  WILLIAMS, Mark Philip
;                  MOLONEY, Margaret Bridget
;                  Holland
;                  EDWARDS, Stirling John
;       TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
;       NUMBER OF SEQUENCES: 50
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: FOLEY & LARDNER
;            STREET: 3000 K Street, N.W.
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: U.S.A.
;            ZIP: 20007-5109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/000,094
;            FILING DATE: 21-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: WO PCT/AU96/00473
;            FILING DATE: 26-JUL-1996
;            APPLICATION NUMBER: AU PN 4439/95
;            FILING DATE: 27-JUL-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BENT, Stephen A.
;            REGISTRATION NUMBER: 29,768
;            REFERENCE/DOCKET NUMBER: 017227/0137
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 672-5300
;            TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 21:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1128 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..1125
;       SEQUENCE DESCRIPTION: SEQ ID NO: 21:
US-09-000-094-21

  Query Match             100.0%;  Score 20;  DB 4;  Length 1128;
  Best Local Similarity   100.0%;  Pred. No. 0.082;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 6
US-09-000-094-23
; Sequence 23, Application US/09000094
; Patent No. 6365160
;   GENERAL INFORMATION:
;       APPLICANT: WEBB, Elizabeth Ann
;                  MARGETTS, Mary Brigid
;                  COX, John Cooper
;                  FRAZER, Ian
;                  MCMILLAN, Nigel Alan John
;                  WILLIAMS, Mark Philip
;                  MOLONEY, Margaret Bridget
;                  Holland
;                  EDWARDS, Stirling John
;       TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
;       NUMBER OF SEQUENCES: 50
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: FOLEY & LARDNER
;            STREET: 3000 K Street, N.W.
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: U.S.A.
;            ZIP: 20007-5109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/000,094
;            FILING DATE: 21-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: WO PCT/AU96/00473
;            FILING DATE: 26-JUL-1996
;            APPLICATION NUMBER: AU PN 4439/95
;            FILING DATE: 27-JUL-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BENT, Stephen A.
;            REGISTRATION NUMBER: 29,768
;            REFERENCE/DOCKET NUMBER: 017227/0137
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 672-5300
;            TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 23:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1398 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
```

```
;           NAME/KEY:  CDS
;           LOCATION:  1..1395
;       SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-000-094-23

  Query Match            100.0%;  Score 20;  DB 4;  Length 1398;
  Best Local Similarity  100.0%;  Pred. No. 0.084;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 7
US-09-000-094-45
; Sequence 45, Application US/09000094
; Patent No. 6365160
;  GENERAL INFORMATION:
;       APPLICANT: WEBB, Elizabeth Ann
;                  MARGETTS, Mary Brigid
;                  COX, John Cooper
;                  FRAZER, Ian
;                  MCMILLAN, Nigel Alan John
;                  WILLIAMS, Mark Philip
;                  MOLONEY, Margaret Bridget
;                  Holland
;                  EDWARDS, Stirling John
;       TITLE OF INVENTION: PAPILLOMAVIRUS POLYPROTEIN CONSTRUCTS
;       NUMBER OF SEQUENCES: 50
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: FOLEY & LARDNER
;            STREET: 3000 K Street, N.W.
;            CITY: Washington
;            STATE: D.C.
;            COUNTRY: U.S.A.
;            ZIP: 20007-5109
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/000,094
;            FILING DATE: 21-Apr-1998
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;          · APPLICATION NUMBER: WO PCT/AU96/00473
;            FILING DATE: 26-JUL-1996
;            APPLICATION NUMBER: AU PN 4439/95
;            FILING DATE: 27-JUL-1995
;       ATTORNEY/AGENT INFORMATION:
;            NAME: BENT, Stephen A.
;            REGISTRATION NUMBER: 29,768
;            REFERENCE/DOCKET NUMBER: 017227/0137
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (202) 672-5300
;            TELEFAX: (202) 672-5399
;   INFORMATION FOR SEQ ID NO: 45:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 4770 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  1..4761
;       SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-09-000-094-45

  Query Match            100.0%;  Score 20;  DB 4;  Length 4770;
  Best Local Similarity  100.0%;  Pred. No. 0.099;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db       2357 TGCTAATTCGGTGCTACCTG 2376


RESULT 8
US-09-521-526-2
; Sequence 2, Application US/09521526
; Patent No. 6290965
;  GENERAL INFORMATION:
;    APPLICANT:  JANSEN, KATHRIN U.
;    APPLICANT:  HOFMANN, KATHRYN J.
;    TITLE OF INVENTION:  DNA ENCODING HUMAN PAPILLOMA VIRUS TYPE
;    TITLE OF INVENTION:  6A
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTINE E. CARTY
;      STREET:  126 E. LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/521,526
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/310,468
;      FILING DATE:  22-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CARTY, CHRISTINE E.
;      REGISTRATION NUMBER:  36,099
;      REFERENCE/DOCKET NUMBER:  19307
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-6734
;      TELEFAX:  (908) 594-4720
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8010 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-521-526-2

  Query Match            100.0%;  Score 20;  DB 3;  Length 8010;
  Best Local Similarity  100.0%;  Pred. No. 0.11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


RESULT 9
PCT-US95-11859-2
; Sequence 2, Application PC/TUS9511859
;  GENERAL INFORMATION:
;    APPLICANT:  JANSEN, KATHRIN U.
;    APPLICANT:  HOFMANN, KATHRYN J.
;    TITLE OF INVENTION:  DNA ENCODING HUMAN PAPILLOMA VIRUS TYPE
;    TITLE OF INVENTION:  6A
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CHRISTINE E. CARTY
;      STREET:  126 E. LINCOLN AVENUE - P.O. BOX 2000
```

```
;       CITY:  RAHWAY
;       STATE:  NJ
;       COUNTRY:  USA
;       ZIP:  07065-0907
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/11859
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/310,468
;       FILING DATE:  22-SEP-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  CARTY, CHRISTINE E.
;       REGISTRATION NUMBER:  36,099
;       REFERENCE/DOCKET NUMBER:  19307 PCT
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-6734
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  8010 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
PCT-US95-11859-2

  Query Match             100.0%;  Score 20;  DB 5;  Length 8010;
  Best Local Similarity   100.0%;  Pred. No. 0.11;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


RESULT 10
US-09-228-986-5/c
; Sequence 5, Application US/09228986
; Patent No. 6359198
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT: Nieuwenhuizen, Niels
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling
;  FILE REFERENCE: 11000/1020
;  CURRENT APPLICATION NUMBER: US/09/228,986
;  CURRENT FILING DATE: 1999-01-12
;  NUMBER OF SEQ ID NOS: 130
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 3340
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-09-228-986-5

  Query Match             77.0%;  Score 15.4;  DB 4;  Length 3340;
  Best Local Similarity   94.1%;  Pred. No. 31;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACC 18
              ||||||||| |||||||
Db       1176 GCTAATTCGTTGCTACC 1160


RESULT 11
US-09-345-236B-76
; Sequence 76, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
;  APPLICANT: Becnel, James J.
;  APPLICANT:  Tukuo, Fukuda
;  APPLICANT:  Moser, Bettina
;  APPLICANT:  Cockburn, Andrew
;  APPLICANT:  White, Susan E.
;  APPLICANT:  Undeen, Albert H.
;  TITLE OF INVENTION: No. 6521454el Baculoviruses,  Insecticidal
;  TITLE OF INVENTION:  Compositions, and Methods for Control of Invertebrates
;  FILE REFERENCE: 21042.0004
;  CURRENT APPLICATION NUMBER: US/09/345,236B
;  CURRENT FILING DATE:  1999-06-30
;  NUMBER OF SEQ ID NOS: 148
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 76
;   LENGTH: 692
;   TYPE: DNA
;   ORGANISM: mosquito baculovirus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (27)...(689)
US-09-345-236B-76

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 692;
  Best Local Similarity   88.9%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 CTAATTCGGTGCTACCTG 20
              ||||||||| ||| ||||
Db        165 CTAATTCGGCGCTGCCTG 182


RESULT 12
US-09-345-236B-78
; Sequence 78, Application US/09345236B
; Patent No. 6521454
; GENERAL INFORMATION:
;  APPLICANT: Becnel, James J.
;  APPLICANT:  Tukuo, Fukuda
;  APPLICANT:  Moser, Bettina
;  APPLICANT:  Cockburn, Andrew
;  APPLICANT:  White, Susan E.
;  APPLICANT:  Undeen, Albert H.
;  TITLE OF INVENTION: No. 6521454el Baculoviruses,  Insecticidal
;  TITLE OF INVENTION:  Compositions, and Methods for Control of Invertebrates
;  FILE REFERENCE: 21042.0004
;  CURRENT APPLICATION NUMBER: US/09/345,236B
;  CURRENT FILING DATE:  1999-06-30
;  NUMBER OF SEQ ID NOS: 148
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 78
;   LENGTH: 692
;   TYPE: DNA
;   ORGANISM: mosquito baculovirus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (13)...(165)
US-09-345-236B-78

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 692;
  Best Local Similarity   88.9%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 CTAATTCGGTGCTACCTG 20
              ||||||||| ||| ||||
Db        165 CTAATTCGGCGCTGCCTG 182


RESULT 13
US-09-345-236B-80/c
; Sequence 80, Application US/09345236B
; Patent No. 6521454
```

```
; GENERAL INFORMATION:
;  APPLICANT: Becnel, James J.
;  APPLICANT:  Tukuo, Fukuda
;  APPLICANT:  Moser, Bettina
;  APPLICANT:  Cockburn, Andrew
;  APPLICANT:  White, Susan E.
;  APPLICANT:  Undeen, Albert H.
;  TITLE OF INVENTION: No. 6521454el Baculoviruses,  Insecticidal
;  TITLE OF INVENTION:  Compositions, and Methods for Control of Invertebrates
;  FILE REFERENCE: 21042.0004
;  CURRENT APPLICATION NUMBER: US/09/345,236B
;  CURRENT FILING DATE:  1999-06-30
;  NUMBER OF SEQ ID NOS: 148
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 80
;   LENGTH: 692
;   TYPE: DNA
;   ORGANISM: mosquito baculovirus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)...(366)
;   NAME/KEY: CDS
;   LOCATION: (529)...(690)
US-09-345-236B-80

  Query Match             74.0%;  Score 14.8;  DB 4;  Length 692;
  Best Local Similarity   88.9%;  Pred. No. 54;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 CTAATTCGGTGCTACCTG 20
              |||||||||| ||| ||||
Db        528 CTAATTCGGCGCTGCCTG 511


RESULT 14
US-09-149-476-80
; Sequence 80, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: 186 Human Secreted proteins
;  FILE REFERENCE: PZ002P1
;  CURRENT APPLICATION NUMBER: US/09/149,476
;  CURRENT FILING DATE: 1998-09-08
;  EARLIER APPLICATION NUMBER: PCT/US98/04493
;  EARLIER FILING DATE: 1998-03-06
;  EARLIER APPLICATION NUMBER: 60/040,162
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,333
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/038,621
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,626
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,334
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,336
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/040,163
;  EARLIER FILING DATE: 1997-03-07
;  EARLIER APPLICATION NUMBER: 60/047,600
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,615
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,597
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,502
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,633
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,583
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,617
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,618
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,503
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,592
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,581
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,584
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,500
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,587
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,492
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,598
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,613
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,582
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,596
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,612
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,632
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/047,601
;  EARLIER FILING DATE: 1997-05-23
;  EARLIER APPLICATION NUMBER: 60/043,580
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,568
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,314
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,569
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,311
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,671
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,674
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,669
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,312
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,313
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,672
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/043,315
;  EARLIER FILING DATE: 1997-04-11
;  EARLIER APPLICATION NUMBER: 60/048,974
;  EARLIER FILING DATE: 1997-06-06
;  EARLIER APPLICATION NUMBER: 60/056,886
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,877
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,889
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,893
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,630
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,878
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,662
;  EARLIER FILING DATE: 1997-08-22
;  EARLIER APPLICATION NUMBER: 60/056,872
;  EARLIER FILING DATE: 1997-08-22
```

; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,874
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,910
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,864
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,631
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,845
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,892
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match 74.0%; Score 14.8; DB 4; Length 1285;
Best Local Similarity 88.9%; Pred. No. 58;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 CTAATTCGGTGCTACCTG 20
              ||||||| ||||| ||||
Db        419 CTAATTCTGTGCTGCCTG 436
```

RESULT 15
US-09-345-217-1
; Sequence 1, Application US/09345217
; Patent No. 6268142
; GENERAL INFORMATION:
; APPLICANT: DUFF, GORDON W.
; APPLICANT: COX, ANGELA
; APPLICANT: CAMP, NICOLA J.
; APPLICANT: DIGIOVINE, FRANCESCO S.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR DISEASES ASSOCIATED
; TITLE OF INVENTION: WITH AN IL-1 INFLAMMATORY HAPLOTYPE
; FILE REFERENCE: MSA-010.02
; CURRENT APPLICATION NUMBER: US/09/345,217
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: PCT/GB98/01481
; EARLIER FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 9711040.7
; EARLIER FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 11970
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-345-217-1

Query Match 74.0%; Score 14.8; DB 3; Length 11970;
Best Local Similarity 88.9%; Pred. No. 78;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 CTAATTCGGTGCTACCTG 20
              ||||||| || |||||||
Db       5364 CTAATTCAGTTCTACCTG 5381
```

Search completed: August 23, 2003, 12:15:10
Job time : 52 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:14:58 ; Search time 1848 Seconds
(without alignments)
263.035 Million cell updates/sec

Title: US-09-927-585A-2
Perfect score: 20
Sequence: 1 tgctaattcggtgctacctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	17.4	87.0	887	29	CC149256	CC149256 ZMMBBb15H
2	16.8	84.0	370	10	BE436705	BE436705 EST407823
3	16.8	84.0	528	28	AQ137540	AQ137540 HS_2252_A
4	16.8	84.0	1096	29	CNS04H07	AL290320 Tetraodon
c 5	16.4	82.0	600	29	CC122092	CC122092 NDL.46E23
c 6	16.4	82.0	854	29	CC134974	CC134974 NDL.91K13
c 7	16.4	82.0	932	29	CC136978	CC136978 NDL.91B22
8	15.8	79.0	226	10	BB570973	BB570973 BB570973
c 9	15.8	79.0	273	28	BH878622	BH878622 hs83c03.b
c 10	15.8	79.0	300	9	AU114943	AU114943 AU114943
c 11	15.8	79.0	376	28	BH397091	BH397091 AG-ND-168
c 12	15.8	79.0	400	29	CC159313	CC159313 ig01c01.b
13	15.8	79.0	427	10	BF937261	BF937261 fm62g01.y
14	15.8	79.0	470	28	AQ678663	AQ678663 HS_2094_B
c 15	15.8	79.0	523	13	BQ378925	BQ378925 RC2-UT008
16	15.8	79.0	532	12	BJ058697	BJ058697 BJ058697
c 17	15.8	79.0	536	14	CD437384	CD437384 EL01N0372
c 18	15.8	79.0	540	28	AQ766138	AQ766138 HS_5477_B
19	15.8	79.0	543	12	BJ037290	BJ037290 BJ037290
20	15.8	79.0	545	14	CB081561	CB081561 hk53e02.g
21	15.8	79.0	551	12	BI984033	BI984033 fu07e12.y
22	15.8	79.0	575	28	AQ491873	AQ491873 RPCI-11-2
23	15.8	79.0	577	12	BM192517	BM192517 dai18e07.
24	15.8	79.0	586	14	CB879626	CB879626 HP03A18w
25	15.8	79.0	598	12	BJ071409	BJ071409 BJ071409
c 26	15.8	79.0	635	28	BZ067787	BZ067787 ljq96g12.
c 27	15.8	79.0	646	28	AZ082092	AZ082092 RPCI-23-2
c 28	15.8	79.0	670	29	CC089529	CC089529 CSU-K33r.
c 29	15.8	79.0	710	2	HSM086542	Bx493714 Homo sapi
30	15.8	79.0	740	29	CC129477	CC129477 NDL.76P12
c 31	15.8	79.0	743	12	BM407709	BM407709 EST582036
32	15.8	79.0	782	29	BZ698014	BZ698014 PUBMZ53TD
33	15.8	79.0	786	29	BZ698353	BZ698353 PUCDX35TD
34	15.8	79.0	825	29	BZ677257	BZ677257 PUBJU13TD
c 35	15.8	79.0	826	10	BE884306	BE884306 601505775
36	15.8	79.0	855	29	CC066386	CC066386 CSU-K33r.
37	15.8	79.0	869	13	BQ734750	BQ734750 AGENCOURT
38	15.8	79.0	890	13	BQ730599	BQ730599 AGENCOURT
c 39	15.8	79.0	925	29	CC079842	CC079842 CSU-K33r.
c 40	15.8	79.0	948	29	CNS02NBR	AL205200 Tetraodon
41	15.8	79.0	1052	29	CC224089	CC224089 CH261-41G
c 42	15.8	79.0	1228	29	CC232880	CC232880 CH261-49K
c 43	15.4	77.0	237	9	AI335576	AI335576 qt30h03.x
c 44	15.4	77.0	238	13	BQ345302	BQ345302 MR4-NT014
c 45	15.4	77.0	304	9	AA188347	AA188347 zp72c07.s

ALIGNMENTS

RESULT 1
CC149256
LOCUS CC149256 887 bp DNA linear GSS 24-APR-2003
DEFINITION ZMMBBb15H20.r ZMMBBb Zea mays subsp. mays genomic clone ZMMBBb15H20 3', genomic survey sequence.
ACCESSION CC149256
VERSION CC149256.1 GI:30094449
KEYWORDS GSS.
SOURCE Zea mays subsp. mays (maize)
ORGANISM Zea mays subsp. mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 887)
AUTHORS Yu,Y., Kim,H.R., Hatfield,J., Soderlund,C., Bharti,A.K., Messing,J. and Wing,R.
TITLE Sequencing of the maize genome
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR PRimers

```
          FORWARD: T7
          BACKWARD: M13r
          Plate: 15  row: H  column: 20
          Seq primer: M13r
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .887
                     /organism="Zea mays subsp. mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4578"
                     /clone="ZMMBBb15H20"
                     /lab_host="DH10B"
                     /clone_lib="ZMMBBb"
                     /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                     HindIII; Zea mays L. ssp. mays"
BASE COUNT      277 a    145 c    178 g    286 t      1 others
ORIGIN

  Query Match            87.0%;  Score 17.4;  DB 29;  Length 887;
  Best Local Similarity   94.7%;  Pred. No. 3.8e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              |||||||| ||||||||||
Db        645 TGCTAATTGGGTGCTACCT 663


RESULT 2
BE436705
LOCUS       BE436705                370 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST407823 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
            clone cLEG33L22, mRNA sequence.
ACCESSION   BE436705
VERSION     BE436705.1  GI:9434548
KEYWORDS    EST.
SOURCE      Lycopersicon esculentum (tomato)
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamiids; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
            Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
            Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
            ,S.D.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage
  JOURNAL   Unpublished
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Lycopersicon esculentum"
                     /mol_type="mRNA"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG33L22"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /clone_lib="tomato breaker fruit, TIGR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; Fruit were harvested at the breaker stage
                     (first sign of lycopene accumulation on the blossom end of
                     the fruit).  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
BASE COUNT      107 a     77 c     79 g    107 t
ORIGIN

  Query Match            84.0%;  Score 16.8;  DB 10;  Length 370;
  Best Local Similarity   90.0%;  Pred. No. 5.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              || ||||| ||||||||||
Db        317 TGTTAATTTGGTGCTACCTG 336


RESULT 3
AQ137540
LOCUS       AQ137540                528 bp    DNA     linear   GSS 23-SEP-1998
DEFINITION  HS_2252_A2_B06_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2252 Col=12 Row=C, genomic survey
            sequence.
ACCESSION   AQ137540
VERSION     AQ137540.1  GI:3524210
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 528)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2252  row: C  column: 12
            Class: BAC ends
            High quality sequence stop: 528.
FEATURES             Location/Qualifiers
     source          1. .528
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2252 Col=12 Row=C"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      135 a    150 c    114 g    126 t      3 others
ORIGIN

  Query Match            84.0%;  Score 16.8;  DB 28;  Length 528;
  Best Local Similarity   90.0%;  Pred. No. 6.4e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              |||| | ||||||||||||||
Db        369 TGCTGACTCGGTGCTACCTG 388


RESULT 4
CNS04H07
LOCUS       CNS04H07                1096 bp    DNA     linear   GSS 01-SEP-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            109D19 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL290320
VERSION     AL290320.1  GI:8028900
KEYWORDS    GSS; genome survey sequence.
```

```
SOURCE      Tetraodon nigroviridis
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetradontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 1096)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .1096
                     /organism="Tetraodon nigroviridis"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:99883"
                     /clone="109D19"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG109CB10SP1~end :
                     PUC-Ori"
BASE COUNT      310 a    232 c    232 g    316 t      6 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 29;  Length 1096;
  Best Local Similarity   90.0%;  Pred. No. 8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              |||||||| ||||||| |||
Db        724 TGCTAATTAGGTGCTAGCTG 743


RESULT 5
CC122092/c
LOCUS       CC122092                600 bp    DNA     linear   GSS 16-APR-2003
DEFINITION  NDL.46E23.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
            NDL.46E23, genomic survey sequence.
ACCESSION   CC122092
VERSION     CC122092.1  GI:29991147
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished
COMMENT     Other_GSSs: NDL.46E23.T7
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .600
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="Liverpool"
                     /db_xref="taxon:7159"
                     /clone="NDL.46E23"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
BASE COUNT      180 a    160 c    181 g     79 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 29;  Length 600;
  Best Local Similarity   94.4%;  Pred. No. 1.1e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACC 18
              ||||| ||||||||||||
Db        349 TGCTATTTCGGTGCTACC 332


RESULT 6
CC134974/c
LOCUS       CC134974                854 bp    DNA     linear   GSS 16-APR-2003
DEFINITION  NDL.91K13.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
            NDL.91K13, genomic survey sequence.
ACCESSION   CC134974
VERSION     CC134974.1  GI:30004029
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
REFERENCE   1  (bases 1 to 854)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished
COMMENT     Other_GSSs: NDL.91K13.T7
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .854
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="Liverpool"
                     /db_xref="taxon:7159"
                     /clone="NDL.91K13"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
BASE COUNT      239 a    208 c    240 g    167 t
```

```
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 29;  Length 854;
  Best Local Similarity   94.4%;  Pred. No. 1.2e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACC 18
              ||||| ||||||||||||
Db        399 TGCTATTTCGGTGCTACC 382


RESULT 7
CC136978/c
LOCUS       CC136978                932 bp    DNA     linear   GSS 16-APR-2003
DEFINITION  NDL.91B22.SP6 Notre Dame Liverpool Aedes aegypti genomic clone
            NDL.91B22, genomic survey sequence.
ACCESSION   CC136978
VERSION     CC136978.1  GI:30006033
KEYWORDS    GSS.
SOURCE      Aedes aegypti (yellow fever mosquito)
  ORGANISM  Aedes aegypti
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   Loftus,B., Shetty,J., Knudson,D. and Severson,D.
  TITLE     BAC end sequencing of Aedes aegypti
  JOURNAL   Unpublished
COMMENT     Other_GSSs: NDL.91B22.T7
            Contact: Brendan Loftus
            Department of Eukaryotic Genomics
            TIGR
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-838-3543
            Fax: 301-838-0208
            Email: enta@tigr.org
            Library was provided by David Severson
            Seq primer: SP6
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1. .932
                     /organism="Aedes aegypti"
                     /mol_type="genomic DNA"
                     /strain="Liverpool"
                     /db_xref="taxon:7159"
                     /clone="NDL.91B22"
                     /clone_lib="Notre Dame Liverpool"
                     /note="Vector: pECBAC1; Site_1: Hind III; The library was
                     prepared from whole body tissue of newly hatched L1 larvae
                     by David Severson at the University of Notre Dame and
                     Hongbin Zhang"
BASE COUNT      262 a    221 c    265 g    184 t
ORIGIN

  Query Match             82.0%;  Score 16.4;  DB 29;  Length 932;
  Best Local Similarity   94.4%;  Pred. No. 1.2e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACC 18
              ||||| ||||||||||||
Db        396 TGCTATTTCGGTGCTACC 379


RESULT 8
BB570973
LOCUS       BB570973                226 bp    mRNA    linear   EST 29-NOV-2000
DEFINITION  BB570973 RIKEN full-length enriched, 0 day neonate skin Mus
            musculus cDNA clone 4633401D12 5', mRNA sequence.
ACCESSION   BB570973
VERSION     BB570973.1  GI:11461881
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 226)
  AUTHORS   Aizawa,K., Akahira,S., Akimura,T., Arai,A., Arakawa,T., Carninci,P.
            , Hanagaki,T., Hayatsu,N., Hiraoka,T., Hirozane,T., Hodoyama,Y.,
            Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kawai,J., Kojima,Y., Konno
            ,H., Kusakabe,M., Matsuyama,T., Miyazaki,A., Nakamura,M., Nishi,K.,
            Nomura,K., Numazaki,R., Okazaki,Y., Okido,T., Owa,C., Sakai,C.,
            Sakai,K., Sasaki,D., Sato,K., Shibata,K., Shibata,Y., Shinagawa,A.,
            Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Tanaka
            ,T., Toya,T., Watahiki,A., Yamamura,T., Yasunishi,A., Yoshida,K.,
            Yoshiki,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Aizawa,K. et al. 2000)
  JOURNAL   Unpublished
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki
            ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Thermostabilization and thermoactivation of thermolabile enzymes by
            trehalose and its application for the synthesis of full length
            cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
             Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
            Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
            ,Y. and Hayashizaki,Y.
             Automated filtration-based high-throughput plasmid preparation
            system. Genome Res. 9 (5), 463-470 (1999)
             Carninci,P. and Hayashizaki,Y.
             High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
            19-44 (1999)
             Please visit our web site (http://genome.rtc.riken.go.jp) for
            further details.
FEATURES             Location/Qualifiers
     source          1. .226
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="4633401D12"
                     /sex="mixed"
                     /tissue_type="skin"
                     /dev_stage="0 day neonate"
                     /lab_host="DH10B"
                     /clone_lib="RIKEN full-length enriched, 0 day neonate
                     skin"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3']. cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 100.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cloned into the XhoI and BamHI sites.
                     Vector: a modified pBluescript KS(+) after bulk excision
                     from Lambda FLC I"
BASE COUNT       63 a     56 c     52 g     55 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 10;  Length 226;
  Best Local Similarity   89.5%;  Pred. No. 1.6e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          1 TGCTAATTCGGTGCTACCT 19
              | |||||| ||||||||||
Db        174 TTCTAATTAGGTGCTACCT 192


RESULT 9
BH878622/c
LOCUS       BH878622                273 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  hs83c03.b1 WGS-ZmaysF (JM107 adapted methyl filtered) Zea mays
            genomic clone hs83c03 5', genomic survey sequence.
ACCESSION   BH878622
VERSION     BH878622.1  GI:22114519
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 273)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Zea mays (methyl-filtered)
  JOURNAL   Unpublished
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: hs83  row: c  column: 03
            Seq primer: -21M13UnivFwd
            Class: shotgun
            High quality sequence stop: 273.
FEATURES             Location/Qualifiers
     source          1. .273
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="hs83c03"
                     /lab_host="JM107 or DH5a"
                     /clone_lib="WGS-ZmaysF (JM107 adapted methyl filtered)"
                     /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
                     The vector was digested with XbaI and one nucleotide was
                     added by  fill in in the recessive 3' end. The genomic DNA
                     was nebulized, end  repaired, adaptor ligated and size
                     fractionated using sephadex. The  resulting fragments were
                     between 0.8 and 3 kb and were cloned into  the vector
                     (.x/y reads in M13mp19, .b/g reads in pUC19). The same
                     ligation was transformed in either JM107 or DH5a. "
BASE COUNT       78 a     58 c     47 g     90 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 28;  Length 273;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              || ||||| ||||||||||
Db         44 GCAAATTCAGTGCTACCTG 26


RESULT 10
AU114943/c
LOCUS       AU114943                300 bp    mRNA    linear   EST 19-OCT-2000
DEFINITION  AU114943 unpublished oligo-capped cDNA library Caenorhabditis
            elegans cDNA clone yk723h1 3', mRNA sequence.
ACCESSION   AU114943
VERSION     AU114943.1  GI:10928510
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 300)
  AUTHORS   Kohara,Y., Shin-i,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
  TITLE     A complementary view of the C. elegans genome
  JOURNAL   Unpublished
COMMENT     Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .300
                     /organism="Caenorhabditis elegans"
                     /mol_type="mRNA"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone="yk723h1"
                     /sex="Hermaphrodite"
                     /tissue_type="whole animal"
                     /dev_stage="varied"
                     /clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT       95 a     62 c     46 g     97 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 300;
  Best Local Similarity   89.5%;  Pred. No. 1.7e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              | |||||||||||||| ||
Db        261 TACTAATTCGGTGCTAACT 243


RESULT 11
BH397091/c
LOCUS       BH397091                376 bp    DNA     linear   GSS 11-DEC-2001
DEFINITION  AG-ND-168H24.TF ND-TAM Anopheles gambiae genomic clone AG-ND-168H24
            , genomic survey sequence.
ACCESSION   BH397091
VERSION     BH397091.1  GI:17343307
KEYWORDS    GSS.
SOURCE      Anopheles gambiae (African malaria mosquito)
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
            Anopheles.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Hong,Y.S., Hogan,J.R., Wang,X., Sarkar,A., Sim,C., Loftus,B.J., Ren
            ,C., Huff,E.R., Carlile,J.L., Black,K., Zhang,H.-B., Gardner,M.J.
            and Collins,F.H.
  TITLE     Construction of a BAC library and generation of BAC end
            sequence-tagged connectors for genome sequencing of the African
            malaria mosquito Anopheles gambiae
  JOURNAL   Mol. Genet. Genomics 268 (6), 720-728 (2003)
  MEDLINE   22542063
   PUBMED   12655398
COMMENT     Other_GSSs: AG-ND-168H24.TR
            Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            This clone is from an A. gambiae BAC library (ND-TAM) provided by
            F.H. Collins and sequenced by The Institute for Genomic Research
            (TIGR). The BAC library was generated from A. gambiae PEST strain
```

DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M Universty BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.
Seq primer: M13 For
Class: BAC ends.

```
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Anopheles gambiae"
                     /mol_type="genomic DNA"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="AG-ND-168H24"
                     /clone_lib="ND-TAM"
                     /note="Vector: pECBAC1; Site_1: HindIII"
BASE COUNT      108 a     75 c    101 g     92 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 28;  Length 376;
  Best Local Similarity   89.5%;  Pred. No. 1.8e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| |||| ||||||||
Db        248 TGCTATTTCGCTGCTACCT 230


RESULT 12
CC159313/c
LOCUS       CC159313               400 bp    DNA     linear   GSS 29-APR-2003
DEFINITION  ig01c01.b1 WGS-ZmaysF (DH5a methyl filtered) Zea mays genomic clone
            ig01c01, genomic survey sequence.
ACCESSION   CC159313
VERSION     CC159313.1  GI:30184091
KEYWORDS    GSS.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Zea mays (methyl-filtered)
  JOURNAL   Unpublished
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: ig01  row: c  column: 01
            Seq primer: -21M13UnivFwd
            Class: shotgun
            High quality sequence stop: 400.
FEATURES             Location/Qualifiers
     source          1. .400
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ig01c01"
                     /lab_host="DH5a"
                     /clone_lib="WGS-ZmaysF (DH5a methyl filtered)"
                     /note="Organ: immature ears; Site_1: Xba I; Site_2: Xba I;
                     The vector was digested with XbaI and one nucleotide was
                     added by fill in in the recessive 3' end. The genomic DNA
                     was nebulized, end repaired, adaptor ligated and size
                     fractionated using sephadex. The resulting fragments were
                     between 0.8 and 3 kb and were cloned into the vector (.x/y
                     reads in M13mp19, .b/g reads in pUC19). The same ligation
                     was transformed into DH5a."
BASE COUNT      128 a     73 c     72 g    127 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 29;  Length 400;
  Best Local Similarity   89.5%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              || ||||| ||||||||||
Db        210 GCAAATTCAGTGCTACCTG 192


RESULT 13
BF937261
LOCUS       BF937261               427 bp    mRNA    linear   EST 13-FEB-2002
DEFINITION  fm62g01.y1 Zebrafish adult retina cDNA Danio rerio cDNA clone
            IMAGE:4200001 5', mRNA sequence.
ACCESSION   BF937261
VERSION     BF937261.1  GI:12354581
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 427)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Library constructed by: Chandra Tucker and Gregory Niemi DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution: RessourcenZentrumPrimarDatenbank, Berlin, Germany
            (web address: www.rzpd.de)
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 421.
FEATURES             Location/Qualifiers
     source          1. .427
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="wild-type"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:4200001"
                     /sex="mixed"
                     /dev_stage="1-2 years"
                     /lab_host="E.Coli XL1-Blue MRF' (XL1-Blue MRF')"
                     /clone_lib="Zebrafish adult retina cDNA"
                     /note="Vector: Lambda ZAP II (pBluescript SK-); Site_1:
                     EcoRI; Site_2: SalI; This Zebrafish library was
                     constructed by Dr. Susan E. Brockerhoff (email:
                     sbrocker@u.washington.edu) RZPD library number: 760"
BASE COUNT      135 a    106 c     97 g     89 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 10;  Length 427;
  Best Local Similarity   89.5%;  Pred. No. 1.9e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              |||||||| |||||| |||
```

```
Db        310 GCTAATTCAGTGCTAGCTG 328


RESULT 14
AQ678663
LOCUS       AQ678663                 470 bp    DNA     linear   GSS 25-JUN-1999
DEFINITION  HS_2094_B2_F10_MR CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2094 Col=20 Row=L, genomic survey
            sequence.
ACCESSION   AQ678663
VERSION     AQ678663.1  GI:5227467
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htsc.washington.edu
            Plate: 2094  row: L  column: 20
            Seq primer: M13 Reverse
            Class: BAC ends
            High quality sequence stop: 470.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2094 Col=20 Row=L"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      123 a    100 c     79 g    168 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 28;  Length 470;
  Best Local Similarity  89.5%;  Pred. No. 2e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              |||| |||| |||||||||
Db        252 TGCTGATTCAGTGCTACCT 270


RESULT 15
BQ378925/c
LOCUS       BQ378925                 523 bp    mRNA    linear   EST 21-MAY-2002
DEFINITION  RC2-UT0087-200900-110-h04 UT0087 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BQ378925
VERSION     BQ378925.1  GI:21054439
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 523)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC2&t2=RC2-UT0087-
            200900-110-h04&t3=2000-09-20&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 21
            High quality sequence stop: 523.
FEATURES             Location/Qualifiers
     source          1. .523
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="UT0087"
                     /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
BASE COUNT      163 a    109 c     99 g    152 t
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 13;  Length 523;
  Best Local Similarity  89.5%;  Pred. No. 2e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              || |||||| |||||||||
Db        448 GCAAATTCGCTGCTACCTG 430


Search completed: August 23, 2003, 12:13:24
Job time : 1852 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:11:44 ; Search time 141 Seconds
(without alignments)
318.888 Million cell updates/sec

Title: US-09-927-585A-2
Perfect score: 20
Sequence: 1 tgctaattcggtgctacctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	17	85.0	427	13	US-10-027-632-74539	Sequence 74539, A
2	15.8	79.0	281	9	US-09-764-878-249	Sequence 249, App
3	15.8	79.0	281	9	US-09-764-860-928	Sequence 928, App
4	15.8	79.0	281	14	US-10-079-854-249	Sequence 249, App
5	15.8	79.0	281	14	US-10-074-095-928	Sequence 928, App
6	15.8	79.0	287	9	US-09-764-878-36	Sequence 36, Appl
7	15.8	79.0	287	9	US-09-764-860-170	Sequence 170, App
8	15.8	79.0	287	14	US-10-079-854-36	Sequence 36, Appl
9	15.8	79.0	287	14	US-10-074-095-170	Sequence 170, App
c 10	15.8	79.0	4682	14	US-10-128-714-257	Sequence 257, App
c 11	15.8	79.0	14468	12	US-10-295-362-19	Sequence 19, Appl
c 12	15.4	77.0	3340	14	US-10-101-464A-5	Sequence 5, Appli
13	15.2	76.0	245	13	US-10-044-090-752	Sequence 752, App
c 14	15.2	76.0	457	13	US-10-027-632-35375	Sequence 35375, A
15	15.2	76.0	457	13	US-10-027-632-60413	Sequence 60413, A
16	15.2	76.0	457	13	US-10-027-632-298306	Sequence 298306,
c 17	15.2	76.0	611	13	US-10-027-632-198643	Sequence 198643,
c 18	15.2	76.0	611	13	US-10-027-632-198644	Sequence 198644,
19	15.2	76.0	795	13	US-10-027-632-165367	Sequence 165367,
20	15	75.0	675	13	US-10-027-632-262303	Sequence 262303,
21	14.8	74.0	451	11	US-09-918-995-12873	Sequence 12873, A
22	14.8	74.0	654	13	US-10-027-632-243504	Sequence 243504,
c 23	14.8	74.0	690	13	US-10-027-632-27924	Sequence 27924, A
c 24	14.8	74.0	690	13	US-10-027-632-27925	Sequence 27925, A
c 25	14.8	74.0	690	13	US-10-027-632-27926	Sequence 27926, A
26	14.8	74.0	1158	13	US-10-027-632-216899	Sequence 216899,
27	14.8	74.0	1285	11	US-09-809-391-80	Sequence 80, Appl
28	14.8	74.0	2029	13	US-10-027-632-264047	Sequence 264047,
29	14.8	74.0	3334	13	US-10-108-605-316	Sequence 316, App
c 30	14.8	74.0	7036	14	US-10-178-782-3	Sequence 3, Appli
31	14.8	74.0	8326	11	US-09-764-872-716	Sequence 716, App
32	14.8	74.0	11970	10	US-09-845-129-1	Sequence 1, Appli
33	14.8	74.0	11970	12	US-10-172-919-1	Sequence 1, Appli
34	14.8	74.0	11970	14	US-10-167-127-1	Sequence 1, Appli
35	14.8	74.0	21732	11	US-09-764-872-717	Sequence 717, App
36	14.8	74.0	152331	13	US-10-095-407-16	Sequence 16, Appl
c 37	14.4	72.0	36	11	US-09-848-616-83	Sequence 83, Appl
38	14.4	72.0	650	13	US-10-027-632-202394	Sequence 202394,
39	14.4	72.0	650	13	US-10-027-632-202395	Sequence 202395,
40	14.4	72.0	674	13	US-10-027-632-160032	Sequence 160032,
41	14.4	72.0	674	13	US-10-027-632-160033	Sequence 160033,
c 42	14.4	72.0	1348	13	US-10-027-632-252549	Sequence 252549,
c 43	14.4	72.0	1348	13	US-10-027-632-252550	Sequence 252550,
44	14.4	72.0	4375	10	US-09-965-703-60	Sequence 60, Appl
c 45	14.2	71.0	355	9	US-09-813-358-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1
US-10-027-632-74539
; Sequence 74539, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 74539
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-74539

Query Match 85.0%; Score 17; DB 13; Length 427;
Best Local Similarity 89.5%; Pred. No. 25;
Matches 17; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 GCTAATTCGGTGCTACCTG 20
              |:||||||||||||| |||
Db        276 GYTAATTCGGTGCTATCTG 294
```

```
RESULT 2
US-09-764-878-249
; Sequence 249, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 249
;   LENGTH: 281
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-878-249

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 281;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 3
US-09-764-860-928
; Sequence 928, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008
;  CURRENT APPLICATION NUMBER: US/09/764,860
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1198
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 928
;   LENGTH: 281
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-860-928

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 281;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 4
US-10-079-854-249
; Sequence 249, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 249
;   LENGTH: 281
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-079-854-249

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 281;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 5
US-10-074-095-928
; Sequence 928, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008C1
;  CURRENT APPLICATION NUMBER: US/10/074,095
;  CURRENT FILING DATE:  2002-02-14
;  PRIOR APPLICATION NUMBER: 09/764,860
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
;  PRIOR APPLICATION NUMBER: 60/235,834
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: 60/234,274
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/234,223
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: 60/228,924
;  PRIOR FILING DATE: 2000-08-30
;  PRIOR APPLICATION NUMBER: 60/224,518
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/236,369
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: 60/224,519
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,964
;  PRIOR FILING DATE: 2000-07-26
```

; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 79.0%; Score 15.8; DB 14; Length 281;
Best Local Similarity 89.5%; Pred. No. 1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 6
US-09-764-878-36
; Sequence 36, Application US/09764878
; Patent No. US20020090615A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121
;  CURRENT APPLICATION NUMBER: US/09/764,878
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 287
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-878-36

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 287;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 7
US-09-764-860-170
; Sequence 170, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008
;  CURRENT APPLICATION NUMBER: US/09/764,860
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 1198
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 170
;   LENGTH: 287
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (1)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-860-170

  Query Match             79.0%;  Score 15.8;  DB 9;  Length 287;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 8
US-10-079-854-36
; Sequence 36, Application US/10079854
; Publication No. US20030054368A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PA121C1
;  CURRENT APPLICATION NUMBER: US/10/079,854
;  CURRENT FILING DATE:  2002-02-22
;  Prior Application removed - See File Wrapper or Palm
;  NUMBER OF SEQ ID NOS: 428
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 287
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-079-854-36

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 287;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 9
US-10-074-095-170
; Sequence 170, Application US/10074095
; Publication No. US20030077704A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC008C1
;  CURRENT APPLICATION NUMBER: US/10/074,095
;  CURRENT FILING DATE:  2002-02-14
;  PRIOR APPLICATION NUMBER: 09/764,860
;  PRIOR FILING DATE: 2001-01-17
;  PRIOR APPLICATION NUMBER: 60/179,065
;  PRIOR FILING DATE: 2000-01-31
;  PRIOR APPLICATION NUMBER: 60/180,628
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: 60/214,886
;  PRIOR FILING DATE: 2000-06-28
;  PRIOR APPLICATION NUMBER: 60/217,487
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,758
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/220,963
;  PRIOR FILING DATE: 2000-07-26
;  PRIOR APPLICATION NUMBER: 60/217,496
;  PRIOR FILING DATE: 2000-07-11
;  PRIOR APPLICATION NUMBER: 60/225,447
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/218,290
;  PRIOR FILING DATE: 2000-07-14
;  PRIOR APPLICATION NUMBER: 60/225,757
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/226,868
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: 60/216,647
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,267
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/216,880
;  PRIOR FILING DATE: 2000-07-07
;  PRIOR APPLICATION NUMBER: 60/225,270
;  PRIOR FILING DATE: 2000-08-14
;  PRIOR APPLICATION NUMBER: 60/251,869
;  PRIOR FILING DATE: 2000-12-08
```

; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14

```
;  PRIOR APPLICATION NUMBER: 60/241,808
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,826
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,786
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/241,221
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/246,475
;  PRIOR FILING DATE: 2000-11-08
;  PRIOR APPLICATION NUMBER: 60/231,243
;  PRIOR FILING DATE: 2000-09-08

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 287;
  Best Local Similarity   89.5%;  Pred. No. 1e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCT 19
              ||||| ||| |||||||||
Db        102 TGCTATTTCTGTGCTACCT 120


RESULT 10
US-10-128-714-257/c
; Sequence 257, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 257
;   LENGTH: 4682
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-257

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 4682;
  Best Local Similarity   89.5%;  Pred. No. 1.4e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              || ||||||| ||||||||
Db       4659 GCGAATTCGGCGCTACCTG 4641


RESULT 11
US-10-295-362-19/c
; Sequence 19, Application US/10295362
; Publication No. US20030143745A1
; GENERAL INFORMATION:
;  APPLICANT: Aventis Pharma
;  APPLICANT:  Martinez, Asuncion
;  APPLICANT:  Kolvek, Steven
;  TITLE OF INVENTION: Genetically Modified Bacterial Strains and No. US20030143745A1
;  TITLE OF INVENTION:  Expressing and Assaying Natural Products
;  FILE REFERENCE: USAV2001/0145 PCT
;  CURRENT APPLICATION NUMBER: US/10/295,362
;  CURRENT FILING DATE:  2002-11-15
;  PRIOR APPLICATION NUMBER: US 60/335,181
;  PRIOR FILING DATE: 2001-11-15
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
;   LENGTH: 14468
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: pMBD14
US-10-295-362-19

  Query Match             79.0%;  Score 15.8;  DB 12;  Length 14468;
  Best Local Similarity   89.5%;  Pred. No. 1.5e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACCTG 20
              |||||| |||||||||| |
Db      12215 GCTAATGCGGTGCTACCGG 12197


RESULT 12
US-10-101-464A-5/c
; Sequence 5, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
;  APPLICANT: Strabala, Timothy
;  APPLICANT:  Nieuwenhuizen, Nicolaas
;  APPLICANT:  Higgins, Colleen M.
;  TITLE OF INVENTION: Compositions Isolated from Plant Cells
;  TITLE OF INVENTION:  and Their Use in the Modification of Plant Cell Signaling
;  FILE REFERENCE: 11000.1020c2
;  CURRENT APPLICATION NUMBER: US/10/101,464A
;  CURRENT FILING DATE:  2002-03-18
;  PRIOR APPLICATION NUMBER: 09/704,302
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 09/228,986
;  PRIOR FILING DATE: 1999-01-12
;  PRIOR APPLICATION NUMBER: 60/162,866
;  PRIOR FILING DATE: 1999-11-01
;  PRIOR APPLICATION NUMBER: PCT/US00/00724
;  PRIOR FILING DATE: 2000-01-11
;  NUMBER OF SEQ ID NOS: 989
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 3340
;   TYPE: DNA
;   ORGANISM: Pinus radiata
US-10-101-464A-5

  Query Match             77.0%;  Score 15.4;  DB 14;  Length 3340;
  Best Local Similarity   94.1%;  Pred. No. 2.2e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GCTAATTCGGTGCTACC 18
              ||||||||| |||||||
Db       1176 GCTAATTCGTTGCTACC 1160


RESULT 13
US-10-044-090-752
; Sequence 752, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
;  APPLICANT: Olga Bandman
;  TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
;  FILE REFERENCE: PA-0028 US
;  CURRENT APPLICATION NUMBER: US/10/044,090
```

```
;  CURRENT FILING DATE:  2002-01-09
;  NUMBER OF SEQ ID NOS: 850
;  SOFTWARE: PERL Program
; SEQ ID NO 752
;    LENGTH: 245
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: misc_feature
;    OTHER INFORMATION: Incyte ID No. US20020137081A1 034851.1
US-10-044-090-752

  Query Match             76.0%;  Score 15.2;  DB 13;  Length 245;
  Best Local Similarity   85.0%;  Pred. No. 2.1e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              |||||| ||||||  |||||
Db        121 TGCTAAGTCGGTGAGACCTG 140


RESULT 14
US-10-027-632-35375/c
; Sequence 35375, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 35375
;    LENGTH: 457
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-35375

  Query Match             76.0%;  Score 15.2;  DB 13;  Length 457;
  Best Local Similarity   85.0%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              |||||||||  ||||||  ||
Db         40 TGCTAATTCTCTGCTACATG 21


RESULT 15
US-10-027-632-60413
; Sequence 60413, Application US/10027632
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 60413
;    LENGTH: 457
;    TYPE: DNA
;    ORGANISM: Human
US-10-027-632-60413

  Query Match             76.0%;  Score 15.2;  DB 13;  Length 457;
  Best Local Similarity   85.0%;  Pred. No. 2.3e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              |||||||||  ||||||  ||
Db        418 TGCTAATTCTCTGCTACATG 437


Search completed: August 23, 2003, 13:53:17
Job time : 142 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 08:28:23 ; Search time 993 Seconds
(without alignments)
823.960 Million cell updates/sec

Title: US-09-927-585A-2
Perfect score: 20
Sequence: 1 tgctaattcggtgctacctg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	6	AX301236	AX301236 Sequence
2		20	100.0	20	6	E05232	E05232 Part of DNA
3		20	100.0	20	6	E05256	E05256 Primer for
4	c	20	100.0	30	6	I34500	I34500 Sequence 9
5	c	20	100.0	30	6	I57335	I57335 Sequence 9
6	c	20	100.0	30	6	I73213	I73213 Sequence 9
7		20	100.0	697	14	AF126428	AF126428 Human pap
8		20	100.0	1107	6	AR202651	AR202651 Sequence
9		20	100.0	1128	6	AR202652	AR202652 Sequence
10		20	100.0	1398	6	AR202653	AR202653 Sequence
11		20	100.0	4770	6	AR202671	AR202671 Sequence
12		20	100.0	7902	14	PAPA6B	X00203 Genital hum
13		20	100.0	8010	6	AR169144	AR169144 Sequence
14		20	100.0	8010	14	PPHE6E	L41216 Human papil
15		20	100.0	8012	14	AF092932	AF092932 Human pap
16		16.8	84.0	3891	14	HPU40822	U40822 Human papil
17		16.8	84.0	7887	14	AF436130	AF436130 Human pap
18	c	16.8	84.0	165468	2	AC129311	AC129311 Mus muscu
19	c	16.8	84.0	173316	2	AC141726	AC141726 Apis mell
20	c	16.8	84.0	231561	10	AL731823	AL731823 Mouse DNA
21		16.8	84.0	238650	2	AC135943	AC135943 Rattus no
22	c	16.8	84.0	242958	2	AC112011	AC112011 Rattus no
23	c	16.8	84.0	245636	2	AC095871	AC095871 Rattus no
24		16.8	84.0	252185	2	AC122980	AC122980 Rattus no
25	c	16.8	84.0	290670	2	AC121043	AC121043 Rattus no
26	c	16.8	84.0	348250	1	AP003592	AP003592 Nostoc sp
27	c	16.4	82.0	168268	10	AL807796	AL807796 Mouse DNA
28	c	16.4	82.0	298050	1	AL935260	AL935260 Lactobaci
29	c	16	80.0	147349	2	AP005800	AP005800 Oryza sat
30		16	80.0	192687	2	AC131316	AC131316 Mus muscu
31	c	16	80.0	198067	2	BX530057	BX530057 Danio rer
32	c	16	80.0	213001	2	AC118630	AC118630 Mus muscu
33	c	16	80.0	214621	2	AC124524	AC124524 Mus muscu
34	c	16	80.0	223389	2	AC108412	AC108412 Mus muscu
35		15.8	79.0	466	5	AF255339	AF255339 Xenopus l
36	c	15.8	79.0	667	1	AF246444	AF246444 Acidithio
37		15.8	79.0	1751	7	BPC31ATTP	X57036 Bacteriopha
38		15.8	79.0	2065	5	BC044687	BC044687 Xenopus l
39		15.8	79.0	2443	3	AB017002	AB017002 Dugesia j
40		15.8	79.0	7916	14	HPU31784	U31784 Human papil
41	c	15.8	79.0	28939	9	AL391808	AL391808 Human DNA
42	c	15.8	79.0	34216	3	U57054	U57054 Caenorhabdi
43	c	15.8	79.0	34589	9	AC112240	AC112240 Homo sapi
44	c	15.8	79.0	37482	3	AC024790	AC024790 Caenorhab
45		15.8	79.0	41489	7	BPH6589	AJ006589 Bacteriop

ALIGNMENTS

RESULT 1
AX301236
LOCUS AX301236 20 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 5 from Patent WO0185994.
ACCESSION AX301236
VERSION AX301236.1 GI:17382323
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Perseu,S., de Montis,A. and Floris,M.M.
TITLE Method and means for identifying hpv virus
JOURNAL Patent: WO 0185994-A 5 15-NOV-2001;
Bioanalisi Centro Sud S.N.C. Di Perseu Siniblado EC. (IT)

```
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
                     /note="primer representative of region E6 of strains of
                     HPV virus having low oncogenic hazard"
BASE COUNT        3 a      5 c      5 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 2
E05232
LOCUS       E05232                    20 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Part of DNA sequence of Human papillomavirus 6.
ACCESSION   E05232
VERSION     E05232.1  GI:2173422
KEYWORDS    JP 1993192200-A/2.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.
  TITLE     DETECTION OF HUMAN PAPILLOMA VIRUS
  JOURNAL   Patent: JP 1993192200-A 2 03-AUG-1993;
            TAKARA SHUZO CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            OS   Human papillomavirus
            PN   JP 1993192200-A/2
            PD   03-AUG-1993
            PF   19-AUG-1991 JP 1991230839
            PR   20-AUG-1990 JP 90P   217067
            PI   OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI
            FUKUSHIMA MICHIO,
            PI   FUJINAGA KEI
            PC   C12Q1/70,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      5 c      5 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 3
E05256
LOCUS       E05256                    20 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Primer for amplifying Human papillomavirus gene.
ACCESSION   E05256
VERSION     E05256.1  GI:2173446
KEYWORDS    JP 1993192200-A/26.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Okazawa,K., Shimada,M., Katou,I., Fukushima,M. and Fujinaga,K.
  TITLE     DETECTION OF HUMAN PAPILLOMA VIRUS
  JOURNAL   Patent: JP 1993192200-A 26 03-AUG-1993;
            TAKARA SHUZO CO LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1993192200-A/26
            PD   03-AUG-1993
            PF   19-AUG-1991 JP 1991230839
            PR   20-AUG-1990 JP 90P   217067
            PI   OKAZAWA KAZUHIDE, SHIMADA MASAMITSU, KATOU IKUNOSHIN, PI
            FUKUSHIMA MICHIO,
            PI   FUJINAGA KEI
            PC   C12Q1/70,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      5 c      5 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1e+02;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db          1 TGCTAATTCGGTGCTACCTG 20


RESULT 4
I34500/c
LOCUS       I34500                    30 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 9 from patent US 5597910.
ACCESSION   I34500
VERSION     I34500.1  GI:1825291
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Gudibande,S.R. and Kenten,J.H.
  TITLE     Electrochemiluminescent label for DNA probe assays
  JOURNAL   Patent: US 5597910-A 9 28-JAN-1997;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        9 a      7 c      7 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 5
I57335/c
LOCUS       I57335                    30 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 9 from patent US 5610017.
ACCESSION   I57335
```

```
VERSION     I57335.1  GI:2482399
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Gudibande,S.R. and Kenten,J.H.
  TITLE     Method for conducting a polymerase chain reaction using an improved
            electrochemiluminescent label
  JOURNAL   Patent: US 5610017-A 9 11-MAR-1997;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        9 a      7 c      7 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 6
I73213/c
LOCUS       I73213 -                   30 bp    DNA     linear   PAT 03-APR-1998
DEFINITION  Sequence 9 from patent US 5686244.
ACCESSION   I73213
VERSION     I73213.1  GI:3009352
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Gudibande,S.R. and Kenten,J.H.
  TITLE     Method for detecting a nucleic acid analyte using an improved
            electrochemiluminescent label
  JOURNAL   Patent: US 5686244-A 9 11-NOV-1997;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        9 a      7 c      7 g      7 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 95;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db         29 TGCTAATTCGGTGCTACCTG 10


RESULT 7
AF126428
LOCUS       AF126428                 697 bp    DNA     linear   VRL 09-MAY-1999
DEFINITION  Human papillomavirus type 6 E6 protein gene, complete cds.
ACCESSION   AF126428
VERSION     AF126428.1  GI:4761532
KEYWORDS    .
SOURCE      Human papillomavirus type 6
  ORGANISM  Human papillomavirus type 6
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 697)
  AUTHORS   Ai,W., Toussaint,E. and Roman,A.
  TITLE     CCAAT displacement protein binds to and negatively regulates human
            papillomavirus type 6 E6, E7, and E1 promoters
  JOURNAL   J. Virol. 73 (5), 4220-4229 (1999)
  MEDLINE   99214361
   PUBMED   10196318
REFERENCE   2  (bases 1 to 697)
  AUTHORS   Ai,W. and Roman,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-FEB-1999) Microbiology and Immunology, Indiana
            University School of Medicine, 635 Barnhill Drive, Indianapolis, IN
            46202-5120, USA
FEATURES             Location/Qualifiers
     source          1. .697
                     /organism="Human papillomavirus type 6"
                     /mol_type="genomic DNA"
                     /strain="W50"
                     /db_xref="taxon:31552"
     CDS             127. .579
                     /codon_start=1
                     /product="E6 protein"
                     /protein_id="AAD29398.1"
                     /db_xref="GI:4761533"
                     /translation="MESANASTSATTIDQLCKTFNLSMHTLQINCVFCKNALTTAEIY
                     SYAYKHLKVLFRGGYPYAACACCLEFHGKINQYRHFDYAGYATTVEEETKQDILDVLI
                     RCYLCHKPLCEVEKVKHILTKARFIKLNCTWKGRCLHCWTTCMEDMLP"
BASE COUNT      233 a    144 c    145 g    175 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 697;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        425 TGCTAATTCGGTGCTACCTG 444


RESULT 8
AR202651
LOCUS       AR202651                 1107 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 19 from patent US 6365160.
ACCESSION   AR202651
VERSION     AR202651.1  GI:21498836
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1107)
  AUTHORS   Webb,E.Ann., Margetts,M.Brigid., Cox,J.Cooper., Frazer,I.,
            McMillan,N.Alan.John., Williams,M.Philip.,
            Moloney,M.Bridget.Holland. and Edwards,S.John.
  TITLE     Papillomavirus polyprotein constructs
  JOURNAL   Patent: US 6365160-A 19 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1107
                     /organism="unknown"
BASE COUNT      336 a    269 c    213 g    289 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 1107;
  Best Local Similarity   100.0%;  Pred. No. 38;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 9
AR202652
LOCUS       AR202652                 1128 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 21 from patent US 6365160.
ACCESSION   AR202652
VERSION     AR202652.1  GI:21498838
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
```

```
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   Webb,E.Ann., Margetts,M.Brigid., Cox,J.Cooper., Frazer,I.,
            McMillan,N.Alan.John., Williams,M.Philip.,
            Moloney,M.Bridget.Holland. and Edwards,S.John.
  TITLE     Papillomavirus polyprotein constructs
  JOURNAL   Patent: US 6365160-A 21 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1128
                     /organism="unknown"
BASE COUNT      358 a    275 c    240 g    255 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 6;  Length 1128;
  Best Local Similarity  100.0%;  Pred. No. 38;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 10
AR202653
LOCUS       AR202653               1398 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 23 from patent US 6365160.
ACCESSION   AR202653
VERSION     AR202653.1  GI:21498839
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1398)
  AUTHORS   Webb,E.Ann., Margetts,M.Brigid., Cox,J.Cooper., Frazer,I.,
            McMillan,N.Alan.John., Williams,M.Philip.,
            Moloney,M.Bridget.Holland. and Edwards,S.John.
  TITLE     Papillomavirus polyprotein constructs
  JOURNAL   Patent: US 6365160-A 23 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .1398
                     /organism="unknown"
BASE COUNT      428 a    331 c    283 g    356 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 6;  Length 1398;
  Best Local Similarity  100.0%;  Pred. No. 36;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        299 TGCTAATTCGGTGCTACCTG 318


RESULT 11
AR202671
LOCUS       AR202671               4770 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 45 from patent US 6365160.
ACCESSION   AR202671
VERSION     AR202671.1  GI:21498861
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4770)
  AUTHORS   Webb,E.Ann., Margetts,M.Brigid., Cox,J.Cooper., Frazer,I.,
            McMillan,N.Alan.John., Williams,M.Philip.,
            Moloney,M.Bridget.Holland. and Edwards,S.John.
  TITLE     Papillomavirus polyprotein constructs
  JOURNAL   Patent: US 6365160-A 45 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .4770
                     /organism="unknown"
BASE COUNT     1597 a    890 c   1071 g   1212 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 6;  Length 4770;
  Best Local Similarity  100.0%;  Pred. No. 26;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db       2357 TGCTAATTCGGTGCTACCTG 2376


RESULT 12
PAPA6B
LOCUS       PAPA6B                 7902 bp    DNA     circular VRL 17-JUL-1998
DEFINITION  Genital human papillomavirus type 6b (HPV6b).
ACCESSION   X00203
VERSION     X00203.1  GI:60955
KEYWORDS    E1 gene; E2 gene; E4 gene; E5a gene; E5b gene; E6 gene; E7 gene; L1
            gene; L2 gene; major capsid protein L1; minor capsid protein L2;
            regulatory protein E2; replication protein E1.
SOURCE      Human papillomavirus type 6b
  ORGANISM  Human papillomavirus type 6b
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1
  AUTHORS   Schwarz,E., Durst,M., Demankowski,C., Lattermann,O., Zech,R.,
            Wolfsperger,E., Suhai,S. and zur Hausen,H.
  TITLE     DNA sequence and genome organization of genital human
            papillomavirus type 6b
  JOURNAL   EMBO J. 2 (12), 2341-2348 (1983)
  MEDLINE   84131949
   PUBMED   6321162
COMMENT     There are two types of open reading frames:
            putative early open reading frames (E) and
            putative late open reading frames (L);
            open reading frames L1 and L2 are expected to code for the viral
            structural polypeptides.
FEATURES             Location/Qualifiers
     source          1. .7902
                     /organism="Human papillomavirus type 6b"
                     /mol_type="genomic DNA"
                     /strain="type 6b (HPV6b)"
                     /db_xref="taxon:10600"
     gene            join(64. .67,102. .554)
                     /gene="E6"
     TATA_signal     64. .67
                     /gene="E6"
     CDS             102. .554
                     /gene="E6"
                     /codon_start=1
                     /product="E6 protein"
                     /protein_id="CAA25018.1"
                     /db_xref="GI:60956"
                     /db_xref="SWISS-PROT:P06462"
                     /translation="MESANASTSATTIDQLCKTFNLSMHTLQINCVFCKNALTTAEIY
                     SYAYKHLKVLFRGGYPYAACACCLEFHGKINQYRHFDYAGYATTVEEETKQDILDVLI
                     RCYLCHKPLCEVEKVKHILTKARFIKLNCTWKGRCLHCWTTCMEDMLP"
     gene            530. .826
                     /gene="E7"
     CDS             530. .826
                     /gene="E7"
                     /codon_start=1
                     /product="E7 protein"
                     /protein_id="CAA25019.1"
                     /db_xref="GI:60957"
                     /db_xref="SWISS-PROT:P06464"
                     /translation="MHGRHVTLKDIVLDLQPPDPVGLHCYEQLVDSSEDEVDEVDGQD
                     SQPLKQHFQIVTCCCGCDSNVRLVVQCTETDIREVQQLLLGTLNIVCPICAPKT"
     gene            832. .2781
                     /gene="E1"
     CDS             832. .2781
                     /gene="E1"
                     /codon_start=1
                     /product="replication protein E1"
```

```
                /protein_id="CAA25020.1"
                /db_xref="GI:60958"
                /db_xref="SWISS-PROT:P03113"
                /db_xref="SWISS-PROT:P06459"
                /translation="MADDSGTENEGSGCTGWFMVEAIVQHPTGTQISDDEDEEVEDSG
                YDMVDFIDDSNITHNSLEAQALFNRQEADTHYATVQDLKRKYLGSPYVSPINTIAEAV
                ESEISPRLDAIKLTRQPKKVKRRLFQTRELTDSGYGYSEVEAGTGTQVEKHGVPENGG
                DGQEKDTGRDIEGEEHTEAEAPTNSVREHAGTAGILELLKCKDLRAALLGKFKECFGL
                SFIDLIRPFKSDKTTCLDWVVAGFGIHHSISEAFQKLIEPLSLYAHIQWLTNAWGMVL
                LVLLRFKVNKSRSTVARTLATLLNIPENQMLIEPPKIQSGVAALYWFRTGISNASTVI
                GEAPEWITRQTVIEHGLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                FLNSNMQAKYVKDCATMCRHYKHAEMRKMSIKQWIKHRGSKIEGTGNWKPIVQFLRHQ
                NIEFIPFLTKFKLWLHGTPKKNCIAIVGPPDTGKSYFCMSLISFLGGTVISHVNSSSH
                FWLQPLVDAKVALLDDATQPCWIYMDTYMRNLLDGNPMSIDRKHKALTLIKCPPLLVT
                SNIDITKEDKYKYLHTRVTTFTFPNPFPFDRNGNAVYELSNTNWKCFFERLSSSLDIQ
                DSEDEEDGSNSQAFRCVPGTVVRTL"
gene            2723. .3829
                /gene="E2"
CDS             2723. .3829
                /gene="E2"
                /codon_start=1
                /product="regulatory protein E2"
                /protein_id="CAA25021.1"
                /db_xref="GI:60959"
                /db_xref="SWISS-PROT:P03119"
                /translation="MEAIAKRLDACQEQLLELYEENSTDLHKHVLHWKCMRHESVLLY
                KAKQMGLSHIGMQVVPPLKVSEAKGHNAIEMQMHLESLLRTEYSMEPWTLQETSYEMW
                QTPPKRCFKKRGKTVEVKFDGCANNTMDYVVWTDVYVQDNDTWVKVHSMVDAKGIYYT
                CGQFKTYYVNFVKEAEKYGSTKHWEVCYGSTVICSPASVSSTTQEVSIPESTTYTPAQ
                TSTLVSSSTKEDAVQTPPRKRARGVQQSPCNALCVAHIGPVDSGNHNLITNNHDQHQR
                RNNSNSSATPIVQFQGESNCLKCFRYRLNDRHRHLFDLISSTWHWASSKAPHKHAIVT
                VTYDSEEQRQQFLDVVKIPPTISHKLGFMSLHLL"
gene            3255. .3584
                /gene="E4"
CDS             3255. .3584
                /gene="E4"
                /codon_start=1
                /product="E4 protein"
                /protein_id="CAA25022.1"
                /db_xref="GI:60960"
                /db_xref="SWISS-PROT:P06459"
                /translation="MGAPNIGKYVMAAQLYVLLHLYLALHKKYPFLNLLHTPPHRPPP
                LCPQAPRKTQCKRRLGNEHEESNSPLATPCVWPTLDPWTVETTTSSLTITTSTKDGTT
                VTVQLRL"
gene            3887. .4162
                /gene="E5a"
CDS             3887. .4162
                /gene="E5a"
                /codon_start=1
                /product="E5a protein"
                /protein_id="CAA25023.1"
                /db_xref="GI:60961"
                /db_xref="SWISS-PROT:P06460"
                /translation="MEVVPVQIAAGTTSTFILPVIIAFVVCFVSIILIVWISEFIVYT
                SVLVLTLLLYLLLWLLLTTPLQFFLLTLLVCYCPALYIHYYIVTTQQ"
gene            4159. .4377
                /gene="E5b"
CDS             4159. .4377
                /gene="E5b"
                /codon_start=1
                /product="E5b protein"
                /protein_id="CAA25024.1"
                /db_xref="GI:60962"
                /db_xref="SWISS-PROT:P06461"
                /translation="MMLTCQFNDGDTWLGLWLLCAFIVGMLGLLLMHYRAVQGDKHTK
                CKKCNKHNCNDDYVTMHYTTDGDYIYMN"
gene            4423. .5802
                /gene="L2"
CDS             4423. .5802
                /gene="L2"
                /codon_start=1
                /product="minor capsid protein L2"
                /protein_id="CAA25025.1"
                /db_xref="GI:60963"
                /db_xref="SWISS-PROT:P03106"
                /translation="MAHSRARRRKRASATQLYQTCKLTGTCPPDVIPKVEHNTIADQI
                LKWGSLGVFFGGLGIGTGSGTGGRTGYVPLQTSAKPSITSGPMARPPVVVEPVAPSDP
                SIVSLIEESAIINAGAPEIVPPAHGGFTITSSETTTPAILDVSVTSHTTTSIFRNPVF
                TEPSVTQPQPPVEANGHILISAPTVTSHPIEEIPLDTFVVSSSDSGPTSSTPVPGTAP
                RPRVGLYSRALHQVQVTDPAFLSTPQRLITYDNPVYEGEDVSVQFSHDSIHNAPDEAF
                MDIIRLHRPAIASRRGLVRYSRIGQRGSMHTRSGKHIGARIHYFYDISPIAQAAEEIE
                MHPLVAAQDDTFDIYAESFEPGINPTQHPVTNISDTYLTSTPNTVTQPWGNTTVPLSL
                PNDLFLQSGPDITFPTAPMGTPFSPVTPALPTGPVFITGSGFYLHPAWYFARKRRKRI
                PLFFSDVAA"
      polyA_signal    4554
                /gene="L2"
      gene            5789. .7291
                /gene="L1"
      CDS             5789. .7291
                /gene="L1"
                /codon_start=1
                /product="major capsid protein L1"
                /protein_id="CAA25026.1"
                /db_xref="GI:60964"
                /db_xref="SWISS-PROT:P03100"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVTRTNIFYHASSSRLLAV
                GHPYFSIKRANKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPFLNKYDDVENSGSGGNPGQDNRVNVGMDYKQTQLCMVGCAPP
                LGEHWGKGKQCTNTPVQAGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVPI
                DICGTTCKYPDYLQMAADPYGDRLFFFLRKEQMFARHFFNRAGEVGEPVPDTLIIKGS
                GNRTSVGSSIYVNTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNQLFVTVVDTTR
                STNMTLCASVTTSSTYTNSDYKEYMRHVEEYDLQFIFQLCSITLSAEVMAYIHTMNPS
                VLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKPDPYKNLSFWEVNLKEKF
                SSELDQYPLGRKFLLQSGYRGRSSIRTGVKRPAVSKASAAPKRKRAKTKR"
      repeat_region   7292. .7315
                /rpt_type=TANDEM
      repeat_region   7316. .7339
                /rpt_type=TANDEM
      polyA_signal    7407
      repeat_region   7681. .7705
                /rpt_type=INVERTED
BASE COUNT     2438 a   1530 c   1699 g   2235 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 14;  Length 7902;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        400 TGCTAATTCGGTGCTACCTG 419


RESULT 13
AR169144
LOCUS       AR169144               8010 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 2 from patent US 6290965.
ACCESSION   AR169144
VERSION     AR169144.1  GI:17906915
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 8010)
  AUTHORS   Jansen,K.U. and Hofmann,K.J.
  TITLE     DNA encoding human papillomavirus type 6A
  JOURNAL   Patent: US 6290965-A 2 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .8010
                     /organism="unknown"
BASE COUNT     2470 a   1527 c   1718 g   2295 t
ORIGIN

  Query Match             100.0%;  Score 20;  DB 6;  Length 8010;
  Best Local Similarity   100.0%;  Pred. No. 23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


RESULT 14
PPHE6E
LOCUS       PPHE6E                  8010 bp    DNA     circular VRL 08-AUG-1995
DEFINITION  Human papillomavirus, type 6a, complete genome.
ACCESSION   L41216
VERSION     L41216.1  GI:940299
KEYWORDS    E1 gene; E2 gene; E4 gene; E5 gene; E6 gene; E7 gene; L1 gene; L2
            gene; complete genome.
SOURCE      Human papillomavirus type 6a
  ORGANISM  Human papillomavirus type 6a
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 8010)
  AUTHORS   Hofmann,K.J., Cook,J.C., Joyce,J.G., Brown,D.R., Schultz,L.D.,
            George,H.A., Rosolowsky,M., Fife,K.H. and Jansen,K.U.
  TITLE     Sequence determination of human papillomavirus type 6a and assembly
            of virus-like particles in Saccharomyces cerevisiae
  JOURNAL   Virology 209 (2), 506-518 (1995)
  MEDLINE   95297152
   PUBMED   7778283
COMMENT     On Aug 10, 1995 this sequence version replaced gi:790893.
            Original source text: Human papillomavirus type 6a DNA.
FEATURES             Location/Qualifiers
     source          1. .8010
                     /organism="Human papillomavirus type 6a"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:37122"
     gene            103. .555
                     /gene="E6"
     CDS             103. .555
                     /gene="E6"
                     /note="base 31 is first base after a stop codon; position
                     103. .105 is first start codon; putative"
                     /codon_start=1
                     /protein_id="AAA74211.1"
                     /db_xref="GI:790894"
                     /translation="MESANASTSATTIDQLCKTFNLSMHTLQINCVFCKNALTTAEIY
                     SYAYKQLKVLFRGGYPYAACACCLEFHGKINQYRHFDYAGYATTVEEETKQDILDVLI
                     RCYLCHKPLCEVEKVKHILTKARFIKLNCTWKGRCLHCWTTCMEDMLP"
     gene            531. .827
                     /gene="E7"
     CDS             531. .827
                     /gene="E7"
                     /note="base 441 is first base after a stop codon; position
                     531. .533 is first start codon; putative"
                     /codon_start=1
                     /protein_id="AAA74212.1"
                     /db_xref="GI:790895"
                     /translation="MHGRHVTLKDIVLDLQPPDPVGLHCYEQLVDSSEDEVDEVDGQD
                     SQPLKQHFQIVTCCCGCDSNVRLVVQCTETDIREVQQLLLGTLDIVCPICAPKT"
     gene            833. .2782
                     /gene="E1"
     CDS             833. .2782
                     /gene="E1"
                     /note="base 716 is first base after a stop codon; position
                     833. .835 is first start codon; putative"
                     /codon_start=1
                     /protein_id="AAA74213.1"
                     /db_xref="GI:790896"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVQHPTGTQISDDEDEEVEDSG
                     YDMVDFIDDSNITHNSLEAQALFNRQEADTHYATVQDLKRKYLGSPYVSPINTIAEAV
                     ESEISPRLDAIKLTRQPKKVKRRLFQTRELTDSGYGYSEVEAGTGTQVEKHGVPENGG
                     DGQEKDTGRDIEGEEHTEAEAPTNSVREHAGTAGILELLKCKDLRAALLGKFKECFGL
                     SFIDLIRPFKSDKTTCADWVVAGFGIHHSISEAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLVRFKVNKSRSTVARTLATLLNIPDNQMLIEPPKIQSGVAALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHGLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCATMCRHYKHAEMRKMSIKQWIKHRGSKIEGTGNWKPIVQFLRHQ
                     NIEFIPFLSKFKLWLHGTPKKNCIAIVGPPDTGKSYFCMSLISFLGGTVISHVNSSSH
                     FWLQPLVDAKVALLDDATQPCWIYMDTYMRNLLDGNPMSIDRKHKALTLIKCPPLLVT
                     SNIDITKEEKYKYLHTRVTTFTFPNPFPFDRNGNAVYELSNANWKCFFERLSSSLDIQ
                     DSEDEEDGSNSQAFRCVPGTVVRTL"
     gene            2724. .3830
                     /gene="E2"
     CDS             2724. .3830
                     /gene="E2"
                     /note="base 2697 is first base after a stop codon;
                     position 2724. .2726 is first start codon.; putative"
                     /codon_start=1
                     /protein_id="AAA74214.1"
                     /db_xref="GI:790897"
                     /translation="MEAIAKRLDACQEQLLELYEENSTDLNKHVLHWKCMRHESVLLY
                     KAKQMGLSHIGMQVVPPLKVSEAKGHNAIEMQMHLESLLKTEYSMEPWTLQETSYEMW
                     QTPPKRCFKKRGKTVEVKFDGCANNTMDYVVWTDVYVQDTDSWVKVHSMVDAKGIYYT
                     CGQFKTYYVNFVKEAEKYGSTKQWEVCYGSTVICSPASVSSTTQEVSIPESTTYTPAQ
                     TSTPVSSSTQEDAVQTPPRKRARGVQQSPCNALCVAHIGPVDSGNHNLITNNHDQHQR
                     RNNSNSSATPIVQFQGESNCLKCFRYRLNDKHRHLFDLISSTWHWASPKAPHKHAIVT
                     VTYHSEEQRQQFLNVVKIPPTIRHKLGFMSLHLL"
     gene            3286. .3585
                     /gene="E4"
     CDS             3286. .3585
                     /gene="E4"
                     /note="base 3241 is first base after a stop codon;
                     position 3286. .3288 is first start codon; putative"
                     /codon_start=1
                     /protein_id="AAA74215.1"
                     /db_xref="GI:790898"
                     /translation="MAAQLYVLLHLYLALHKKYPFLNLLHTPPHRPPPLCPQAPRKTQ
                     CKRRLENEHEESNSHLATPCVWPTLDPWTVETTTSSLTITTSTKEGTTVTVQLRL"
     gene            3888. .4163
                     /gene="E5"
     CDS             3888. .4163
                     /gene="E5"
                     /codon_start=1
                     /protein_id="AAA74216.1"
                     /db_xref="GI:790899"
                     /translation="MEVVPVQIAAGTTSTLILPVIIAFVVCFVSIILIVWISDFIVYT
                     SVLVLTLLLYLLLWLLLTTPLQFFLLTLLVCYCPALYIHHYIVNTQQ"
     gene            4424. .5803
                     /gene="L2"
     CDS             4424. .5803
                     /gene="L2"
                     /note="base 4379 is first base after a stop codon;
                     position 4424. .4426 is first start codon; putative"
                     /codon_start=1
                     /protein_id="AAA74217.1"
                     /db_xref="GI:790900"
                     /translation="MAHSRARRRKRASATQLYQTCKLTGTCPPDVIPKVEHNTIADQI
                     LKWGSLGVFFGGLGIGTGSGTGGRTGYVPLGTSAKPSITSGPMARPPVVVEPVAPSDP
                     SIVSLIEESAIINAGAPEIVPPAHGGFTITSSETTTPAILDVSVTSHTTTSIFRNPVF
                     TEPSVTQPQPPVEANGHILISAPTITSHPIEEIPLDTFVISSSDSGPTSSTPVPGTAP
                     RPRVGLYSRALHQVQVTDPAFLSTPQRLITYDNPVYEGEDVSVQFSHDSIHNAPDEAF
                     MDIIRLHRPAIASRRGLVRYSRIGQRGSMHTRSGKHIGARIHYFYDISPIAQAAEEIE
                     MHPLVAAQDDTFDIYAESFEPDINPTQHPVTNISDTYLTSTPNTVTQPWGNTTVPLSI
                     PNDLFLQSGPDITFPTAPMGTPFSPVTPALPTGPVFITGSGFYLHPAWYFARKRRKRI
                     PLFFSDVAA"
     gene            5790. .7292
                     /gene="L1"
     CDS             5790. .7292
                     /gene="L1"
                     /note="base 5679 is first base after a stop codon;
                     position 5790. .5792 is first start codon.; putative"
                     /codon_start=1
                     /protein_id="AAA74218.1"
                     /db_xref="GI:790901"
                     /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVTRTNIFYHASSSRLLAV
                     GHPYFSIKRANKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                     EVGRGQPLGVGVSGHPFLNKYDDVENSGSGGNPGQDNRVNVGMDYKQTQLCMVGCAPP
                     LGEHWGKGKQCTNTPVQAGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVPI
                     DICGTTCKYPDYLQMAADPYGDRLFFFLRKEQMFARHFFNRAGEVGEPVPDTLIIKGS
                     GNRTSVGSSIYVNTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNQLFVTVVDTTR
```

```
                  STNMTLCASVTTSSTYTNSDYKEYMRHVEEYDLQFIFQLCSITLSAEVMAYIHTMNPS
                  VLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKPDPYKNLSFWEVNLKEKF
                  SSELDQYPLGRKFLLQSGYRGRSSIRTGVKRPAVSKASAAPKRKRAKTKR"
BASE COUNT     2470 a   1527 c   1718 g   2295 t
ORIGIN

  Query Match           100.0%;  Score 20;  DB 14;  Length 8010;
  Best Local Similarity  100.0%;  Pred. No. 23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


RESULT 15
AF092932
LOCUS       AF092932               8012 bp    DNA     linear   VRL 14-NOV-2001
DEFINITION  Human papillomavirus type 6, complete genome.
ACCESSION   AF092932
VERSION     AF092932.1  GI:6002612
KEYWORDS    .
SOURCE      Human papillomavirus type 6
  ORGANISM  Human papillomavirus type 6
            Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
            Papillomavirus.
REFERENCE   1  (bases 1 to 8012)
  AUTHORS   Kovelman,R., Bilter,G.K., Roman,A., Brown,D.R. and Barbosa,M.S.
  TITLE     Human papillomavirus type 6: classification of clinical isolates
            and functional analysis of E2 proteins
  JOURNAL   J. Gen. Virol. 80 (Pt 9), 2445-2451 (1999)
  MEDLINE   99429608
   PUBMED   10501500
REFERENCE   2  (bases 1 to 8012)
  AUTHORS   Kovelman,R., Bilter,G.K., Roman,A., Brown,D.R. and Barbosa,M.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-SEP-1998) Virology, Signal Pharmaceuticals, Inc.,
            5555 Oberlin Drive, San Diego, CA 92121, USA
FEATURES             Location/Qualifiers
     source          1. .8012
                     /organism="Human papillomavirus type 6"
                     /mol_type="genomic DNA"
                     /strain="Human papillomavirus type 6vc"
                     /db_xref="taxon:31552"
     gene            103. .555
                     /gene="E6"
     CDS             103. .555
                     /gene="E6"
                     /codon_start=1
                     /product="E6"
                     /protein_id="AAF00064.1"
                     /db_xref="GI:6002618"
                     /translation="MESANASTSATTIDQLCKTFNLSMHTLQINCVFCKNALTTAEIY
                     SYAYKQLKVLFRGGYPYAACACCLEFHGKINQYRHFDYAGYATTVEEETKQDILDVLI
                     RCYLCHKPLCEVEKVKHILTKARFIKLNCTWKGRCLHCWTTCMEDMLP"
     gene            531. .827
                     /gene="E7"
     CDS             531. .827
                     /gene="E7"
                     /codon_start=1
                     /product="E7"
                     /protein_id="AAF00065.1"
                     /db_xref="GI:6002619"
                     /translation="MHGRHVTLKDIVLDLQPPDPVGLHCYEQLVDSSEDEVDEVDGQD
                     SQPLKQHYQIVTCCCGCDSNVRLVVQCTETDIREVQQLLLGTLNIVCPICAPKT"
     gene            833. .2782
                     /gene="E1"
     CDS             833. .2782
                     /gene="E1"
                     /codon_start=1
                     /product="E1"
                     /protein_id="AAF00059.1"
                     /db_xref="GI:6002613"
                     /translation="MADDSGTENEGSGCTGWFMVEAIVQHPTGTQISDDEDEEVEDSG
                     YDMVDFIDDSNITHNSLEAQALFNRQEADTHYATVQDLKRKYLGSPYVSPINTIAEAV
                     ESEISPRLDAIKLTRQPKKVKRRLFQTRELTDSGYGYSEVEAGTGTQVEKHGVPENGG
                     DGQEKDTGRDIEGEEHTEAEAPTNSVREHAGTAGILELLKCKDLRAALLGKFKECFGL
                     SFIDLIRPFKSDKTTCADWVVAGFGIHHSISEAFQKLIEPLSLYAHIQWLTNAWGMVL
                     LVLLRFKVNKSRSTVARTLATLLNIPENQMLIEPPKIQSGVAALYWFRTGISNASTVI
                     GEAPEWITRQTVIEHGLADSQFKLTEMVQWAYDNDICEESEIAFEYAQRGDFDSNARA
                     FLNSNMQAKYVKDCATMCRHYKHAEMRKMSIKQWIKHRGSKIEGTGNWKPIVQFLRHQ
                     NIEFIPFLSKFKLWLHGTPKKNCIAIVGPPDTGKSYFCMSLISFLGGTVISHVNSSSH
                     FWLQPLVDAKVALLDDATQPCWIYMDTYMRNLLDGNPMSIDRKHKALTLIKCPPLLVT
                     SNIDITKEEKYKYLHTRVTTFTFPNPFPFDRNGNAVYELSNANWKCFFERLSSSLDIQ
                     DSEDEEDGSNSQAFRCVPGTVVRTL"
     gene            2724. .3830
                     /gene="E2"
     CDS             2724. .3830
                     /gene="E2"
                     /codon_start=1
                     /product="E2"
                     /protein_id="AAF00060.1"
                     /db_xref="GI:6002614"
                     /translation="MEAIAKRLDACQEQLLELYEENSTDLNKHVLHWKCIRHESVLLY
                     KAKQMGLSHIGMQVVPPLKVSEAKGHNAIEMQMHLESLLKTEYSMEPWTLQETSYEMW
                     QTPPKRCFKKRGKTVEVKFDGCANNTMDYVVWTDVYVQDTDSWVKVHSMVDAKGIYYT
                     CGQFKTYYVNFVKEAEKYGSTKQWEVCYGSTVICSPASVSSTTQEVSIPESTTYTPAQ
                     TSTPVSSSTQEDAVQTPPRKRARGVQQSPCNALCVAHIGPVDSGNHNVITNNHDQHQR
                     RNNSNSSATPIVQFQGESNCLKCFRYRLNDKHRHLFDLISSTWHWASSKAPHKHAIVT
                     VTYHSEEQRQQFLNVVKIPPTIRHKLGFMSMHLL"
     gene            3256. .3585
                     /gene="E4"
     CDS             3256. .3585
                     /gene="E4"
                     /codon_start=1
                     /product="E4"
                     /protein_id="AAF00061.1"
                     /db_xref="GI:6002615"
                     /translation="MGAPNNGKYVMAAQLYVLLHLYLALHKKYPFLNLLHTPPHRPPP
                     LCPQAPRKTQCKRRLENEHEESNSHLATPCVWPTLDPWTVETTTSSLTITTSTKEGTT
                     VTVQLRL"
     gene            3888. .4163
                     /gene="E5A"
     CDS             3888. .4163
                     /gene="E5A"
                     /codon_start=1
                     /product="E5A"
                     /protein_id="AAF00062.1"
                     /db_xref="GI:6002616"
                     /translation="MEVVPVQIAAGTTSTLILPVIIAFVVCFVSIILIVWISDFIVYT
                     SVLVLTLLLYLLLWLLLTTPLQFFLLTLLVCYCPACYIHHYIVNTQQ"
     gene            4160. .4378
                     /gene="E5B"
     CDS             4160. .4378
                     /gene="E5B"
                     /codon_start=1
                     /product="E5B"
                     /protein_id="AAF00063.1"
                     /db_xref="GI:6002617"
                     /translation="MMLTCQFNDGDTWLGLWLLCAFIVGVLGLLLMHYRAVQGDKHTK
                     CNKCNKHNCNDDYVTMHYDTAGDYIYMN"
     gene            4424. .5803
                     /gene="L2"
     CDS             4424. .5803
                     /gene="L2"
                     /codon_start=1
                     /product="L2"
                     /protein_id="AAF00067.1"
                     /db_xref="GI:6002621"
                     /translation="MAHSRARRRKRASATQLYQTCKLTGTCPPDVIPKVEHNTIADQI
                     LKWGSLGVFFGGLGIGTGSGTGGRTGYVPLGTSAKPSITSGPMARPPVVVEPVAPSDP
                     SIVSLIEESAIINAGAPEIVPPAHGGFTITSSETTTPAILDVSVTSHTTTSIFRNPVF
                     TEPSVTQPQPPVEANGHILISAPTITSHPIEEIPLDTFVISSSDSGPTSSTPVPGTAP
                     RPRVGLYSRALHQVQVTDPAFLSTPQRLITYDNPVYEGEDVSVQFSHDSIHNAPDEAF
                     MDIIRLHRPAIASRRGLVRYSRIGQRGSMHTRSGKHIGARIHYFYDISPIAQAAEEIE
                     MHPLVAAQEDTFDIYAESFEPDINPTQHPVTNISDTYLTSTPNTVTQPWGNTTVPLSI
```

```
                PNDLFLQSGPDITFPTAPMGTPFSPVTPALPTGPVFITGSGFYLHPAWYFARKRRKRI
                PLFFSDVAA"
     gene       5790. .7292
                /gene="L1"
     CDS        5790. .7292
                /gene="L1"
                /codon_start=1
                /product="L1"
                /protein_id="AAF00066.1"
                /db_xref="GI:6002620"
                /translation="MWRPSDSTVYVPPPNPVSKVVATDAYVTRTNIFYHASSSRLLAV
                GHPYFSIKRANKTVVPKVSGYQYRVFKVVLPDPNKFALPDSSLFDPTTQRLVWACTGL
                EVGRGQPLGVGVSGHPFLNKYDDVENSGSGGNPGQDNRVNVGMDYKQTQLCMVGCAPP
                LGEHWGKGKQCTNTPVQAGDCPPLELITSVIQDGDMVDTGFGAMNFADLQTNKSDVPI
                YICGTTCKYPDYLQMAADPYGDRLFFFLRKEQMFARHFFNRAGEVGEPVPDTLIIKGS
                GNRTSVGSSIYVNTPSGSLVSSEAQLFNKPYWLQKAQGHNNGICWGNQLFVTVVDTTR
                STNMTLCASVTTSSTYTNSDYKEYMRHVEEYDLQFIFQLCSITLSAEVMAYIHTMNPS
                VLEDWNFGLSPPPNGTLEDTYRYVQSQAITCQKPTPEKEKPDPYKNLSFWEVNLKEKF
                SSELDQYPLGRKFLLQSGYRGRSSIRTGVKRPAVSKASAAPKRKRAKTKR"
BASE COUNT     2478 a   1526 c   1713 g   2295 t
ORIGIN

  Query Match            100.0%;  Score 20;  DB 14;  Length 8012;
  Best Local Similarity  100.0%;  Pred. No. 23;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGCTAATTCGGTGCTACCTG 20
              ||||||||||||||||||||
Db        401 TGCTAATTCGGTGCTACCTG 420


Search completed: August 23, 2003, 11:11:37
Job time : 997 secs
```

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:14:58 ; Search time 1848 Seconds
(without alignments)
263.035 Million cell updates/sec

Title: US-09-927-585A-1
Perfect score: 20
Sequence: 1 tgtcaaaaaccgttgtgtcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	18	90.0	670	12	BM076659	BM076659 TrEST-A05
c	2	17.4	87.0	629	28	AQ291095	AQ291095 nbxb0038C
c	3	17	85.0	831	28	AZ081985	AZ081985 UP_309-16
	4	17	85.0	863	29	CNS009PI	AL054375 Drosophil
	5	16.8	84.0	297	9	AW454500	AW454500 zehl1605
c	6	16.8	84.0	308	14	CA620055	CA620055 wl1n.pk00
c	7	16.8	84.0	370	12	BM103817	BM103817 fv39c12.x
c	8	16.8	84.0	410	10	BE760467	BE760467 an_1726 A
	9	16.8	84.0	425	12	BM201540	BM201540 C0227A12-
	10	16.8	84.0	473	28	AQ228742	AQ228742 HS_2020_A
	11	16.8	84.0	508	9	AI443887	AI443887 sa44f10.y
	12	16.8	84.0	645	14	CA645372	CA645372 wreln.pk0
c	13	16.8	84.0	669	12	BI830009	BI830009 603080272
	14	16.8	84.0	875	13	BU941677	BU941677 AGENCOURT
c	15	16.8	84.0	929	10	BG335280	BG335280 602403794
c	16	16.8	84.0	999	13	BX329000	BX329000 BX329000
c	17	16.4	82.0	322	10	BG405554	BG405554 sac35g01.
c	18	16.4	82.0	328	10	BG405573	BG405573 sac35h09.
c	19	16.4	82.0	364	10	BG406469	BG406469 sac32g02.
c	20	16.4	82.0	374	9	AL817342	AL817342 AL817342
c	21	16.4	82.0	388	10	BG653697	BG653697 sad77f12.
c	22	16.4	82.0	413	10	BG652756	BG652756 sad71f08.
c	23	16.4	82.0	429	10	BG363250	BG363250 sac18b07.
c	24	16.4	82.0	430	10	BG363563	BG363563 sac15c09.
c	25	16.4	82.0	431	10	BG652072	BG652072 sad73g09.
c	26	16.4	82.0	442	10	BG363641	BG363641 sac16d06.
c	27	16.4	82.0	454	10	BG651780	BG651780 sad60h06.
c	28	16.4	82.0	461	10	BE803096	BE803096 sr51g11.y
c	29	16.4	82.0	476	10	BG363283	BG363283 sac18f06.
c	30	16.4	82.0	479	10	BG363409	BG363409 sac20b08.
c	31	16.4	82.0	482	10	BG725786	BG725786 sae42a01.
c	32	16.4	82.0	491	10	BG363306	BG363306 sac19a04.
c	33	16.4	82.0	492	10	BG363341	BG363341 sac20c09.
c	34	16.4	82.0	492	10	BG652372	BG652372 sad66c01.
c	35	16.4	82.0	493	10	BG359585	BG359585 sac26a04.
c	36	16.4	82.0	497	10	BG725368	BG725368 sae36a02.
c	37	16.4	82.0	499	10	BG651803	BG651803 sad61b11.
c	38	16.4	82.0	499	10	BG653953	BG653953 sad63d01.
c	39	16.4	82.0	502	10	BG651372	BG651372 sad83g12.
c	40	16.4	82.0	503	10	BG725471	BG725471 sae37f04.
c	41	16.4	82.0	504	10	BF423738	BF423738 sr39e09.y
c	42	16.4	82.0	505	10	BG363449	BG363449 sac20g06.
c	43	16.4	82.0	508	10	BG406572	BG406572 sac33h09.
c	44	16.4	82.0	508	10	BG651997	BG651997 sad72f12.
c	45	16.4	82.0	511	12	BG789864	BG789864 sae56h06.

ALIGNMENTS

RESULT 1
BM076659
LOCUS BM076659 670 bp mRNA linear EST 05-FEB-2002
DEFINITION TrEST-A0592 TrEST-A Hypocrea jecorina cDNA clone Tr-A0592 5', mRNA sequence.
ACCESSION BM076659
VERSION BM076659.1 GI:18497841
KEYWORDS EST.
SOURCE Hypocrea jecorina
ORGANISM Hypocrea jecorina
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
REFERENCE 1 (bases 1 to 670)
AUTHORS Chambergo,F.S., Bonaccorsi,E.D., Ferreira,A.J.S., Ramos,A.S.P., Ferreira,J.R.Jr., Abrahao-Neto,J., Farah,J.P.S. and El-Dorry,H.
TITLE Elucidation of the metabolic fate of glucose in the filamentous fungus Trichoderma reesei using expressed sequence tag (EST) analysis and cDNA microarrays
JOURNAL J. Biol. Chem. 277 (16), 13983-13988 (2002)
MEDLINE 21950703
PUBMED 11825887
COMMENT Contact: El-Dorry, Hamza
Department of Biochemistry
Institute of Chemistry. University of Sao Paulo
Avenida Professor Lineu Prestes, 748, Sao Paulo, SP, 05508-900, BRASIL
Tel: (55) 11-38183848

```
            Fax: (55) 11-38183848
            Email: dorry@iq.usp.br
            PCR PRimers
            FORWARD: Universal M13 forward primer
            BACKWARD: Universal M13 reverse primer
            Plate: 7  row: B  column: 4
            Seq primer: M13 reverse primer
            High quality sequence stop: 670
            POLYA=No.
FEATURES              Location/Qualifiers
     source           1. .670
                      /organism="Hypocrea jecorina"
                      /mol_type="mRNA"
                      /strain="QM9414 (ATCC26921)"
                      /db_xref="taxon:51453"
                      /clone="Tr-A0592"
                      /sex="Asexual"
                      /tissue_type="Mycelia"
                      /dev_stage="18 hr Glycerol-grown culture"
                      /lab_host="E. coli SOLR cells (kanamycin resistant)"
                      /clone_lib="TrEST-A"
                      /note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:
                      XhoI; anamorph=Trichoderma reesei; Cloned unidirectionally
                      , 5' end of the cDNA cloned into EcoRI site of
                      pBluescript. Primer: Oligo (dT). Average insert size: 1,2
                      kb; Uni-ZAP XR Vector system -5' adaptor sequence:
                      5'GAATTCGGCACGAG3' -3' adaptor sequence:
                      5'CTCGAGTTTTTTTTTTTTTTTTT3'"
BASE COUNT      164 a    184 c    135 g    185 t      2 others
ORIGIN

  Query Match             90.0%;  Score 18;  DB 12;  Length 670;
  Best Local Similarity   100.0%;  Pred. No. 1e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGT 18
              ||||||||||||||||||
Db        425 TGTCAAAAACCGTTGTGT 442


RESULT 2
AQ291095/c
LOCUS       AQ291095                629 bp    DNA     linear   GSS 03-DEC-1998
DEFINITION  nbxb0038C14r CUGI Rice BAC Library Oryza sativa (japonica
            cultivar-group) genomic clone nbxb0038C14r, genomic survey
            sequence.
ACCESSION   AQ291095
VERSION     AQ291095.1  GI:3952385
KEYWORDS    GSS.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 629)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE     A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: GGAAACAGCTATGACCATG
            Class: BAC ends
            High quality sequence stop: 481.
FEATURES              Location/Qualifiers
     source           1. .629
                      /organism="Oryza sativa (japonica cultivar-group)"
                      /mol_type="genomic DNA"
                      /strain="Japonica"
                      /cultivar="Nipponbare"
                      /db_xref="taxon:39947"
                      /clone="nbxb0038C14r"
                      /tissue_type="Leaf"
                      /lab_host="E. coli DH10B"
                      /clone_lib="CUGI Rice BAC Library"
                      /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2:
                      HindIII; Rice is one of two most popular grains in the
                      world. Half of the world  population especially those
                      inhabiting highly populated areas of the  humid tropics
                      and subtropics, rely on rice as their primary source of
                      carbohydrate. Monocotyledonous rice is a diploid plant
                      (2n=24) with a  haploid genome equivalent of 431 Mbp
                      (Arumuganathan and Earle, 1991).  The relatively small
                      genome of rice, three times larger than that of
                      Arabidopsis, makes it suitable for genomic studies. In
                      order to  facilitate positional cloning, physical mapping
                      and genome sequencing  of rice, we have constructed a BAC
                      library from Oryza sativa, Nipponbare  variety. The
                      library contains 36,864 clones with an average insert size
                      of 128.5 Kb providing 10.9 haploid genome equivalents. The
                      deep  coverage allows the isolation a particular sequence
                      with a probability  of 99.9 %. Two high density filters,
                      each containing 18,432 clones  (doubly spotted), represent
                      the whole library for colony screening."
BASE COUNT      223 a    134 c    109 g    162 t      1 others
ORIGIN

  Query Match             87.0%;  Score 17.4;  DB 28;  Length 629;
  Best Local Similarity   94.7%;  Pred. No. 2.1e+02;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTC 19
              | ||||||||||||||||||
Db        262 TTTCAAAAACCGTTGTGTC 244


RESULT 3
AZ081985/c
LOCUS       AZ081985                831 bp    DNA     linear   GSS 28-JAN-2002
DEFINITION  UP_309-16G_SP6 RPCI11 Human Male BAC Library Homo sapiens genomic
            clone 309-16G, genomic survey sequence.
ACCESSION   AZ081985
VERSION     AZ081985.2  GI:18379387
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 831)
  AUTHORS   Cheung,V.G., Dalrymple,H.L., Narasimhan,S., Watts,J., Schuler,G.,
            Raap,A.K., Morley,M. and Bruzel,A.
  TITLE     A resource of mapped human bacterial artificial chromosome clones
  JOURNAL   Genome Res. 9 (10), 989-993 (1999)
  MEDLINE   99455100
   PUBMED   10523527
COMMENT     On Jan 28, 2002 this sequence version replaced gi:7709240.
            Contact: Arcaro MA, Morley M, Burdick J, Cheung VG
            Department of Pediatrics
            University of Pennsylvania
            3516 Civic Center Blvd, ARC 516, Philadelphia, PA 19104, USA
            Tel: 215 590 2664
            Fax: 215 590 3709
            Email: mlennox@mail.med.upenn.edu
            Plate: 309  row: G  column: 16
            Seq primer: SP6
            Class: BAC ends.
FEATURES              Location/Qualifiers
     source           1. .831
                      /organism="Homo sapiens"
                      /mol_type="genomic DNA"
                      /db_xref="taxon:9606"
                      /clone="309-16G"
```

```
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI11 Human Male BAC Library"
                     /note="Vector: pBACe3.6; RPCI11 Human Male BAC Library"
BASE COUNT       226 a    189 c    174 g    242 t
ORIGIN

  Query Match             85.0%;  Score 17;  DB 28;  Length 831;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches   17;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGT 18
              |||||||||||||||||
Db        655 GTCAAAAACCGTTGTGT 639


RESULT 4
CNS009PI
LOCUS       CNS009PI                863 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR20D02 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL054375
VERSION     AL054375.1  GI:4935346
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 863)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .863
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACR20D02"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT       469 a    155 c     91 g     77 t     71 others
ORIGIN

  Query Match             85.0%;  Score 17;  DB 29;  Length 863;
  Best Local Similarity   89.5%;  Pred. No. 3.7e+02;
  Matches   17;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTC 19
              |||||||:|| ||||||||
Db        560 TGTCAAAWACTGTTGTGTC 578


RESULT 5
AW454500
LOCUS       AW454500                297 bp    mRNA    linear   EST 17-FEB-2000
DEFINITION  zehl1605 Zebrafish Embryonic Heart cDNA Library Danio rerio cDNA
            5', mRNA sequence.
ACCESSION   AW454500
VERSION     AW454500.1  GI:6995287
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 297)
  AUTHORS   Ton,C., Mably,J.D., Dempsey,A.A., Hwang,D.M., Fishman,M.C. and Liew
            ,C.C.
  TITLE     Identification and Characterization of Expressed Sequence Tags from
            an Embryonic Zebrafish Heart cDNA Library
  JOURNAL   Unpublished
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            PCR Primers
            FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
            BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'
            Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1. .297
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /dev_stage="embryonic day 3 post-fertilization"
                     /lab_host="E.coli XL1-Blue mrF'"
                     /clone_lib="Zebrafish Embryonic Heart cDNA Library"
                     /note="Organ: heart; Vector: Lambda ZAP Express; Site_1:
                     EcoRI; Site_2: XhoI; mRNA was purified from embryonic
                     zebrafish hearts (3 day  post-fertilization).  cDNA was
                     synthesized using a XhoI-Oligo  dT adaptor-primer.  EcoRI
                     adaptors were ligated, followed by digestion  with XhoI,
                     for directional cloning into pre-digested lambda ZAP
                     Express vector. "
BASE COUNT       101 a     56 c     61 g     79 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 297;
  Best Local Similarity   90.0%;  Pred. No. 3.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              | |||||||| |||||||||
Db        107 TCTCAAAAACGGTTGTGTCC 126


RESULT 6
CA620055/c
LOCUS       CA620055                308 bp    mRNA    linear   EST 23-NOV-2002
DEFINITION  wl1n.pk0052.f8 wl1n Triticum aestivum cDNA clone wl1n.pk0052.f8 5'
            end, mRNA sequence.
ACCESSION   CA620055
VERSION     CA620055.1  GI:25198351
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 308)
  AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
  TITLE     DuPont Wheat cDNA Sequence
  JOURNAL   Unpublished
```

```
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .308
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wl1n.pk0052.f8"
                     /tissue_type="leaf"
                     /clone_lib="wl1n"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum L.) leaf 7 day old seedling
                     , light grown (normalized)"
BASE COUNT       94 a     69 c     44 g    101 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 14;  Length 308;
  Best Local Similarity   90.0%;  Pred. No. 3.9e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||||  |||| |||
Db         30 TGTCAAAAACCATTGTTTCC 11


RESULT 7
BM103817/c
LOCUS       BM103817                370 bp    mRNA    linear   EST 21-NOV-2001
DEFINITION  fv39c12.x1 Sugano SJD adult male Danio rerio cDNA clone
            IMAGE:5410942 3' similar to SW:RPA5_HUMAN O15160 DNA-DIRECTED RNA
            POLYMERASE I 40 KD POLYPEPTIDE ;, mRNA sequence.
ACCESSION   BM103817
VERSION     BM103817.1  GI:17034887
KEYWORDS    EST.
SOURCE      Danio rerio (zebrafish)
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 370)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished
COMMENT     Other_ESTs: fv39c12.y1
            Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution information can be found through the I.M.A.G.E.
            Consortium/LLNL, send email to: info@image.llnl.gov
            Trace considered overall poor quality
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .370
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:5410942"
                     /sex="male"
                     /tissue_type="whole body"
                     /dev_stage="adult"
                     /lab_host="DH10B (phage resistant)"
                     /clone_lib="Sugano SJD adult male"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACCATGTG);
                     Site_2: DraIII (CACTGTGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert. Size selection was performed to exclude fragments
                     <1.5kb. Library constructed and donated by Dr. Sumio
                     Sugano (University of Tokyo Institute of Medical Science).
                     Custom primers for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA. "
BASE COUNT       93 a    114 c     81 g     82 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 370;
  Best Local Similarity   90.0%;  Pred. No. 4e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||||||||| | ||
Db        200 TGTCAAAAACCGTTGGGGCC 181


RESULT 8
BE760467/c
LOCUS       BE760467                410 bp    mRNA    linear   EST 24-OCT-2000
DEFINITION  an_1726 Aspergillus niger, pYES2 (XhoI-EcoRI) Aspergillus niger
            cDNA clone 1726 3', mRNA sequence.
ACCESSION   BE760467
VERSION     BE760467.1  GI:10183104
KEYWORDS    EST.
SOURCE      Aspergillus niger
  ORGANISM  Aspergillus niger
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 410)
  AUTHORS   Tsang,A. and Storms,R.
  TITLE     Aspergillus niger Expressed Sequence Tags
  JOURNAL   Unpublished
COMMENT     Contact: Tsang, A.
            Dept. of Biology
            Concordia University
            1455 deMaisonneuve Blvd. West, Montreal, Quebec, CANADA, H3G 1M8
            Tel: 514-848-3405
            Fax: 514-848-4504
            Email: tsang@vax2.concordia.ca
            PCR PRimers
            BACKWARD: GGCGTGAATGTAAGCGTGACATAAC
            Insert Length: 410   Std Error: 0.00
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .410
                     /organism="Aspergillus niger"
                     /mol_type="mRNA"
                     /db_xref="taxon:5061"
                     /clone="1726"
                     /lab_host="E. coli"
                     /clone_lib="Aspergillus niger, pYES2 (XhoI-EcoRI)"
                     /note="Vector: pYES2; Site_1: Xho-I; Site_2: EcoRI; cDNA
                     was synthesized with ZAP kit (Stratagene) using poly(A)+
                     RNA isolated from Aspergillus niger grown in both complete
                     and minimal media. Synthesis was primed with oligo(dT)
                     primer/XhoI-linker. EcoRI adaptors were later ligated to
                     polished ends. EcoRI-XhoI-digested cDNA was ligated with
                     EcoRI-XhoI-digested pYES2 (Invitrogen Corp). This vector
                     permits expression of our library in yeast. "
```

```
BASE COUNT      103 a    116 c    101 g     81 t      9 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 10;  Length 410;
  Best Local Similarity   90.0%;  Pred. No. 4.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              | |||||||| |||||||||
Db        270 TATCAAAAACAGTTGTGTCC 251


RESULT 9
BM201540
LOCUS       BM201540                425 bp    mRNA    linear   EST 30-JAN-2002
DEFINITION  C0227A12-3 NIA Mouse 7.5-dpc Whole Embryo cDNA Library (Long) Mus
            musculus cDNA clone C0227A12 3', mRNA sequence.
ACCESSION   BM201540
VERSION     BM201540.1  GI:17755294
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 425)
  AUTHORS   Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Tanaka,T., Lim,M.K.,
            Luo,A. and Ko,M.S.H.
  TITLE     Systematic Analyses of NIA Mouse 7.5-dpc Whole Embryo cDNA Library
            (Long)
  JOURNAL   Unpublished
COMMENT     Contact: Dawood B. Dudekula
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            Plate: C0227  row: A  column: 12
            Seq primer: -21M13 Forward
            High quality sequence stop: 425
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .425
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:C0227A12-3"
                     /db_xref="taxon:10090"
                     /clone="C0227A12"
                     /tissue_type="whole embryo including extraembryonic
                     tissues at 7.5-days postcoitum"
                     /dev_stage="7.5-days postcoitum"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 7.5-dpc Whole Embryo cDNA Library
                     (Long)"
                     /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
                     NotI; Mouse cDNA project by the Laboratory of Genetics,
                     National Institute on Aging (NIA), Intramural Research
                     Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
                     a long-transcript enriched cDNA library (Ref. Genome Res.
                     11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
                     extracted from a pool of four embryos at 7.5-days
                     postcoitum. Double-stranded cDNAs were synthesized with an
                     Oligo(dT) primer [Invitrogen:
                     5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT-3'] from
                     7 ug of total RNA, treated with T4 DNA polymerase, and
                     purified by ethanol-precipitation. The cDNAs were ligated
                     to Lone-linker LL-Sal4, purified by phenol/chloroform, and
                     separated from free linkers by Centricon 100. Then, the
                     cDNAs were amplified by long-range high fidelity PCR using
                     Ex Taq polymerase (Takara) with a primer Sal4-S. The
                     products were purified by phenol/chloroform and Centricon
                     100. The cDNAs were digested with SalI and NotI enzymes
                     and cloned into SalI/NotI site of pSPORT1 plasmid vector.
                     The DH10B E. coli host was transformed with the ligation
                     mixture by the standard chemical method. The average
                     insert size is about 2.2 kb. The library was constructed
                     by Yulan Piao (NIA)."
BASE COUNT      144 a     87 c     76 g    118 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 425;
  Best Local Similarity   90.0%;  Pred. No. 4.1e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||| || |||||
Db        234 TGTCAAAAACCCTTTTGTCC 253


RESULT 10
AQ228742
LOCUS       AQ228742                473 bp    DNA     linear   GSS 26-SEP-1998
DEFINITION  HS_2020_A2_F03_T7 CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=2020 Col=6 Row=K, genomic survey
            sequence.
ACCESSION   AQ228742
VERSION     AQ228742.1  GI:3653971
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 473)
  AUTHORS   Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
  TITLE     Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  MEDLINE   99380589
   PUBMED   10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Sequence Tagged Connector
            Plate: 2020  row: K  column: 6
            Class: BAC ends
            High quality sequence stop: 473.
FEATURES             Location/Qualifiers
     source          1. .473
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /clone="Plate=2020 Col=6 Row=K"
                     /sex="male"
                     /clone_lib="CIT Approved Human Genomic Sperm Library D"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT      157 a    101 c     88 g    126 t      1 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 28;  Length 473;
  Best Local Similarity   90.0%;  Pred. No. 4.2e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||| | ||||||
Db        168 TGTCAAAAACCATAGTGTCC 187


RESULT 11
AI443887
LOCUS       AI443887                508 bp    mRNA    linear   EST 28-NOV-2001
```

```
DEFINITION  sa44f10.y1 Gm-c1004 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1004-2180 5' similar to TR:Q96502 Q96502 COL2. ;, mRNA
            sequence.
ACCESSION   AI443887
VERSION     AI443887.1  GI:4306317
KEYWORDS    EST.
SOURCE      Glycine max (soybean)
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
            ; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 508)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Insert Length: 1480   Std Error: 0.00
            Seq primer: -40RP from Gibco
            High quality sequence stop: 393
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .508
                     /organism="Glycine max"
                     /mol_type="mRNA"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1004-2180"
                     /tissue_type="root"
                     /lab_host="XL10-Gold"
                     /clone_lib="Gm-c1004"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; Root cDNA. The mRNA was isolated from entire roots
                     of 8 day old 'Williams' seedlings which were propagated on
                     paper towels with distilled water. Stratagene's cDNA
                     Synthesis Kit (catalog #200401) was used to synthesize the
                     cDNA. First- strand synthesis was performed with 5-methyl
                     dCTP, hence the ligated cDNA is hemimethylated.
                     Stratagene's first-strand synthesis primer was used
                     [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)-18]. After
                     second-strand synthesis, the cDNA ends were 'polished'
                     with clone Pfu DNA polymerase, ligated to EcoRI adapters,
                     and phosphorylated. The XhoI site within the first-strand
                     synthesis primer was restricted by digestion with XhoI;
                     all XhoI sites in the cDNA would be protected by their
                     hemimethylated status. The cDNA constructs were
                     size-fractionated with a 500bp cutoff, using GibcoBRL Life
                     Technologies' cDNA Size Fractionation column. The column
                     eluent was then ligated into Stratagene's pBluescript II
                     XR Predigested vector (pBluescript II SK(+) that had been
                     digested with EcoRI and XhoI, and phosphorylated). Both
                     the white and blue colonies appear to contain recombinant
                     plasmids with cDNA inserts. Blue colonies 9n=15) have been
                     sequenced, and possess putative cDNA inserts. This library
                     was constructed by Dr. Paul Keim & Virginia H. Coryell,
                     Department of Biology, Box5640, Northern Arizona
                     University, Flagstaff, AZ 86011, Phone: 520-523-1078 (Dr.
                     Paul Keim), 520-523-1372 (Virginia H. Coryell), Fax:
                     520-523-7500, email: paul.keim@nau.edu,
                     virginia.coryell@nau.edu"
BASE COUNT       95 a    193 c    118 g    102 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 9;  Length 508;
  Best Local Similarity   90.0%;  Pred. No. 4.3e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||| || |||||||
Db        447 TGTCAAAAAGCGCTGTGTCC 466


RESULT 12
CA645372
LOCUS       CA645372                645 bp    mRNA    linear   EST 23-NOV-2002
DEFINITION  wreln.pk0094.c4 wreln Triticum aestivum cDNA clone wreln.pk0094.c4
            5' end, mRNA sequence.
ACCESSION   CA645372
VERSION     CA645372.1  GI:25223668
KEYWORDS    EST.
SOURCE      Triticum aestivum (bread wheat)
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Tingey,S.V., Powell,W., Wolters,P., Dolan,M., Hainey,C., Yuan,Z.,
            Miao,G., Caraher,N. and Hanafey,M.K.
  TITLE     DuPont Wheat cDNA Sequence
  JOURNAL   Unpublished
COMMENT     Contact: Scott V. Tingey
            Crop Genetics
            E. I. DuPont de Nemours and Company
            1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
            Tel: 302-631-2602
            Fax: 302-631-2607
            Email: Scott.V.Tingey@USA.dupont.com
            Seq primer: M13.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Triticum aestivum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4565"
                     /clone="wreln.pk0094.c4"
                     /tissue_type="root"
                     /clone_lib="wreln"
                     /note="Vector: pBluescript SK+; Site_1: EcoRI; Site_2:
                     XhoI; Wheat (Triticum aestivum L.) root; normalized from
                     wrel library"
BASE COUNT      170 a    149 c    116 g    170 t     40 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 14;  Length 645;
  Best Local Similarity   90.0%;  Pred. No. 4.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| | |||||||||
Db        584 TGTCAAAATCTGTTGTGTCC 603


RESULT 13
BI830009/c
LOCUS       BI830009                669 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603080272F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171968 5',
            mRNA sequence.
ACCESSION   BI830009
VERSION     BI830009.1  GI:15941559
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11427  row: p  column: 17
            High quality sequence stop: 653.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5171968"
                     /tissue_type="medulla"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_119"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source normal medulla from
                     anonymous male age 27.  Library  is oligo-dT primed and
                     directionally cloned (EcoRV site is destroyed  upon
                     cloning).  Average insert size 1.3 kb, insert size range
                     0.9-3 kb.  Library is normalized and enriched for
                     full-length clones and  was constructed by C. Gruber
                     (Invitrogen).  Research Genetics  tracking code 013. Note:
                     this is a NIH_MGC Library."
BASE COUNT      157 a    197 c    189 g    126 t
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 12;  Length 669;
  Best Local Similarity   90.0%;  Pred. No. 4.5e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||| |||||||||| |
Db        545 TGTCAAACACCGTTGTGTTC 526


RESULT 14
BU941677
LOCUS       BU941677                875 bp    mRNA    linear   EST 18-OCT-2002
DEFINITION  AGENCOURT_10554416 NIH_MGC_127 Homo sapiens cDNA clone
            IMAGE:6713986 5', mRNA sequence.
ACCESSION   BU941677
VERSION     BU941677.1  GI:24130496
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 875)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: NCI
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM3025  row: c  column: 10
            High quality sequence stop: 302.
FEATURES             Location/Qualifiers
     source          1. .875
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6713986"
                     /tissue_type="mixed (pool of 40 RNAs)"
                     /lab_host="DH10B (T1-phage-resistant)"
                     /clone_lib="NIH_MGC_127"
                     /note="Vector: pDNR-LIB; Site_1: SfiI (ggccattatggcc);
                     Site_2: SfiI (ggccgcctcggcc); Double-stranded cDNA was
                     prepared from a pool of 40 cell line polyA+ RNAs (bladder
                     - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
                     4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
                     kidnney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%
                     , ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
                     gland - 1.3%, and skin - 2.3%).  5' and 3' adaptors were
                     used in cloning as follows:
                     5'-AAGCAGTGGTATCAACGCAGAGTGGCCATTACGGCCGGG-3' and
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)NN-3'.  Full-length
                     enriched library was constructed using the Clontech
                     Creator SMART kit and size-selected to contain the 1-2 kb
                     size fraction (other fractions present in NIH_MGC_126 and
                     NIH_MGC_128).  Library created in the laboratory of T.
                     Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH_MGC
                     Library."
BASE COUNT      306 a    193 c    193 g    177 t      6 others
ORIGIN

  Query Match             84.0%;  Score 16.8;  DB 13;  Length 875;
  Best Local Similarity   90.0%;  Pred. No. 4.8e+02;
  Matches   18;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              | ||||||||| ||||||||
Db        839 TCTCAAAAACCCTTGTGTCC 858


RESULT 15
BG335280/c
LOCUS       BG335280                929 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602403794F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4541350 5',
            mRNA sequence.
ACCESSION   BG335280
VERSION     BG335280.1  GI:13141718
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1219  row: l  column: 23
            High quality sequence start: 2
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .929
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4541350"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
```

/clone_lib="NIH_MGC_21"
/note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 222 a 254 c 288 g 165 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 929;
Best Local Similarity 90.0%; Pred. No. 4.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||| |||||||| ||||
Db        914 TGTCAAGAACCGTTGGGTCC 895
```

Search completed: August 23, 2003, 12:13:20
Job time : 1851 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 11:11:44 ; Search time 141 Seconds
(without alignments)
318.888 Million cell updates/sec

Title: US-09-927-585A-1
Perfect score: 20
Sequence: 1 tgtcaaaaaccgttgtgtcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1517243 seqs, 1124081882 residues

Total number of hits satisfying chosen parameters: 3034486

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	12	US-10-286-387-8	Sequence 8, Appli
	2	20	100.0	20	14	US-10-292-848-18	Sequence 18, Appl
	3	20	100.0	23	12	US-10-286-387-9	Sequence 9, Appli
	4	20	100.0	23	12	US-10-286-387-10	Sequence 10, Appl
	5	20	100.0	23	12	US-10-286-387-13	Sequence 13, Appl
c	6	15.8	79.0	1409	14	US-10-128-714-2427	Sequence 2427, Ap
c	7	15.8	79.0	1410	14	US-10-128-714-7427	Sequence 7427, Ap
c	8	15.8	79.0	1518	14	US-10-128-714-6427	Sequence 6427, Ap
c	9	15.8	79.0	1521	14	US-10-128-714-1427	Sequence 1427, Ap
c	10	15.8	79.0	3518	14	US-10-128-714-5427	Sequence 5427, Ap
c	11	15.8	79.0	3520	14	US-10-128-714-427	Sequence 427, App
	12	15.2	76.0	20	12	US-10-286-387-15	Sequence 15, Appl
c	13	15.2	76.0	104	11	US-09-941-492-119	Sequence 119, App
c	14	15.2	76.0	138	11	US-09-941-492-122	Sequence 122, App
c	15	15.2	76.0	431	11	US-09-918-995-33830	Sequence 33830, A
	16	15.2	76.0	458	12	US-10-177-390-5	Sequence 5, Appli
	17	15.2	76.0	477	9	US-09-864-761-6139	Sequence 6139, Ap
	18	15.2	76.0	496	10	US-09-783-590-10792	Sequence 10792, A
	19	15.2	76.0	507	13	US-10-027-632-184710	Sequence 184710,
	20	15.2	76.0	525	14	US-10-255-536-71	Sequence 71, Appl
	21	15.2	76.0	618	13	US-10-027-632-191550	Sequence 191550,
c	22	15.2	76.0	621	13	US-10-027-632-269909	Sequence 269909,
	23	15.2	76.0	779	10	US-09-970-477-3	Sequence 3, Appli
	24	15.2	76.0	822	13	US-10-000-903-3	Sequence 3, Appli
	25	15.2	76.0	879	13	US-10-000-903-9	Sequence 9, Appli
	26	15.2	76.0	1116	13	US-10-000-903-5	Sequence 5, Appli
	27	15.2	76.0	1173	13	US-10-000-903-13	Sequence 13, Appl
c	28	15.2	76.0	1537	13	US-10-027-632-263346	Sequence 263346,
c	29	15.2	76.0	2368	14	US-10-133-642-1	Sequence 1, Appli
	30	15.2	76.0	2832	13	US-10-115-178-2	Sequence 2, Appli
	31	15.2	76.0	8115	10	US-09-778-516A-1	Sequence 1, Appli
c	32	15.2	76.0	8115	10	US-09-778-516A-2	Sequence 2, Appli
c	33	15.2	76.0	9025608	14	US-10-156-761-1	Sequence 1, Appli
c	34	15	75.0	692	9	US-09-910-943-270	Sequence 270, App
c	35	15	75.0	758	9	US-09-910-943-358	Sequence 358, App
c	36	15	75.0	860	13	US-10-027-632-8841	Sequence 8841, Ap
c	37	14.8	74.0	338	11	US-09-803-719-1753	Sequence 1753, Ap
	38	14.8	74.0	598	13	US-10-027-632-214264	Sequence 214264,
	39	14.8	74.0	637	13	US-10-027-632-207791	Sequence 207791,
	40	14.8	74.0	637	13	US-10-027-632-207792	Sequence 207792,
c	41	14.8	74.0	710	13	US-10-027-632-25868	Sequence 25868, A
c	42	14.8	74.0	710	13	US-10-027-632-25869	Sequence 25869, A
c	43	14.8	74.0	710	13	US-10-027-632-25870	Sequence 25870, A
	44	14.8	74.0	726	14	US-10-081-051-75	Sequence 75, Appl
	45	14.8	74.0	742	13	US-10-027-632-34048	Sequence 34048, A

ALIGNMENTS

RESULT 1
US-10-286-387-8
; Sequence 8, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
; APPLICANT: Cytyc Corporation
; TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
; FILE REFERENCE: cym-035CP
; CURRENT APPLICATION NUMBER: US/10/286,387
; CURRENT FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-8

Query Match 100.0%; Score 20; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGTCAAAAACCGTTGTGTCC 20
||||||||||||||||||||
Db 1 TGTCAAAAACCGTTGTGTCC 20

RESULT 2
US-10-292-848-18
; Sequence 18, Application US/10292848
; Publication No. US20030119042A1
; GENERAL INFORMATION:
; APPLICANT: ROSADO, Pedro M.
; APPLICANT: LOPEZ, Gemma L.
; APPLICANT: SEIZ, Antonio M.
; APPLICANT: ALBERDI, Maria D.
; TITLE OF INVENTION: Method for preparing stabilized reaction mixtures, which are t

```
;  TITLE OF INVENTION:  partially dried, comprising at least one enzyme, reaction mixtur
;  TITLE OF INVENTION:  containing said mixtures
;  FILE REFERENCE: 9256-15
;  CURRENT APPLICATION NUMBER: US/10/292,848
;  CURRENT FILING DATE:  2002-11-12
;  PRIOR APPLICATION NUMBER: PCT/ES02/00109
;  PRIOR FILING DATE: 2002-3-11
;  PRIOR APPLICATION NUMBER: ES P0100569
;  PRIOR FILING DATE: 2001-3-12
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 2.0
; SEQ ID NO 18
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: VI-3 oligonucleotide primer
US-10-292-848-18

  Query Match             100.0%;  Score 20;  DB 14;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.73;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 3
US-10-286-387-9
; Sequence 9, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
;  APPLICANT: Cytyc Corporation
;  TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
;  FILE REFERENCE: cym-035CP
;  CURRENT APPLICATION NUMBER: US/10/286,387
;  CURRENT FILING DATE:  2003-02-28
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-9

  Query Match             100.0%;  Score 20;  DB 12;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.75;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 4
US-10-286-387-10
; Sequence 10, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
;  APPLICANT: Cytyc Corporation
;  TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
;  FILE REFERENCE: cym-035CP
;  CURRENT APPLICATION NUMBER: US/10/286,387
;  CURRENT FILING DATE:  2003-02-28
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-10

  Query Match             100.0%;  Score 20;  DB 12;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.75;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 5
US-10-286-387-13
; Sequence 13, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
;  APPLICANT: Cytyc Corporation
;  TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
;  FILE REFERENCE: cym-035CP
;  CURRENT APPLICATION NUMBER: US/10/286,387
;  CURRENT FILING DATE:  2003-02-28
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-13

  Query Match             100.0%;  Score 20;  DB 12;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 0.75;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||||||||||||||
Db          1 TGTCAAAAACCGTTGTGTCC 20


RESULT 6
US-10-128-714-2427/c
; Sequence 2427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2427
;   LENGTH: 1409
;   TYPE: DNA
```

```
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-2427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 1409;
  Best Local Similarity   89.5%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | |||||||| ||||||||
Db        769 GCCAAAAACCCTTGTGTCC 751


RESULT 7
US-10-128-714-7427/c
; Sequence 7427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7427
;   LENGTH: 1410
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-7427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 1410;
  Best Local Similarity   89.5%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | |||||||| ||||||||
Db        770 GCCAAAAACCCTTGTGTCC 752


RESULT 8
US-10-128-714-6427/c
; Sequence 6427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6427
;   LENGTH: 1518
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-6427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 1518;
  Best Local Similarity   89.5%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | |||||||| ||||||||
Db        878 GCCAAAAACCCTTGTGTCC 860


RESULT 9
US-10-128-714-1427/c
; Sequence 1427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1427
;   LENGTH: 1521
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-1427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 1521;
  Best Local Similarity   89.5%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | |||||||| ||||||||
Db        878 GCCAAAAACCCTTGTGTCC 860


RESULT 10
US-10-128-714-5427/c
; Sequence 5427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
```

```
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5427
;   LENGTH: 3518
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-5427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 3518;
  Best Local Similarity   89.5%;  Pred. No. 2.6e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | ||||||||  ||||||||
Db       1878 GCCAAAAACCCTTGTGTCC 1860


RESULT 11
US-10-128-714-427/c
; Sequence 427, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
;  APPLICANT: Jiang, Bo
;  APPLICANT:  Hu, Wenqi
;  APPLICANT:  Tishkoff, Daniel
;  APPLICANT:  Zamudio, Carlos
;  APPLICANT:  Eroshkin, Alexey M
;  APPLICANT:  Lemieux, Sebastien M
;  TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
;  TITLE OF INVENTION:  Methods of Use
;  FILE REFERENCE: 10182-018-999
;  CURRENT APPLICATION NUMBER: US/10/128,714
;  CURRENT FILING DATE:  2002-04-23
;  PRIOR APPLICATION NUMBER: US 60/285,697
;  PRIOR FILING DATE: 2001-04-23
;  PRIOR APPLICATION NUMBER: US 60/287,066
;  PRIOR FILING DATE: 2001-04-27
;  PRIOR APPLICATION NUMBER: US 60/295,890
;  PRIOR FILING DATE: 2001-06-05
;  PRIOR APPLICATION NUMBER: US 60/303,899
;  PRIOR FILING DATE: 2001-07-09
;  PRIOR APPLICATION NUMBER: US 60/316,362
;  PRIOR FILING DATE: 2001-08-31
;  NUMBER OF SEQ ID NOS: 8603
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 427
;   LENGTH: 3520
;   TYPE: DNA
;   ORGANISM: Aspergillus fumigatus
US-10-128-714-427

  Query Match             79.0%;  Score 15.8;  DB 14;  Length 3520;
  Best Local Similarity   89.5%;  Pred. No. 2.6e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GTCAAAAACCGTTGTGTCC 20
              | ||||||||  ||||||||
Db       1877 GCCAAAAACCCTTGTGTCC 1859


RESULT 12
US-10-286-387-15
; Sequence 15, Application US/10286387
; Publication No. US20030143529A1
; GENERAL INFORMATION:
;  APPLICANT: Cytyc Corporation
;  TITLE OF INVENTION: Detection and Typing of Human Papillomavirus Using PNA Probes
;  FILE REFERENCE: cym-035CP
;  CURRENT APPLICATION NUMBER: US/10/286,387
;  CURRENT FILING DATE:  2003-02-28
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial
;   FEATURE:
;   OTHER INFORMATION: Forward primer used to amplify region of E6 ORF of HPV
US-10-286-387-15

  Query Match             76.0%;  Score 15.2;  DB 12;  Length 20;
  Best Local Similarity   85.0%;  Pred. No. 2.2e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              || ||||||||  |||||||
Db          1 TGCCAAAAACCACTGTGTCC 20


RESULT 13
US-09-941-492-119/c
; Sequence 119, Application US/09941492
; Publication No. US20030027250A1
; GENERAL INFORMATION:
;  APPLICANT: Mitchell, Lloyd
;  APPLICANT:  Garcia-Blanco, Mariano M.
;  APPLICANT:  Puttaraju, Madaiah
;  APPLICANT:  Mansfield, Gary S.
;  TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN
;  TITLE OF INVENTION:  SPLICEOSOME MEDIATED RNA TRANS-SPLICING
;  FILE REFERENCE: A31304-BAE (072874.0156)
;  CURRENT APPLICATION NUMBER: US/09/941,492
;  CURRENT FILING DATE:  2002-04-01
;  PRIOR APPLICATION NUMBER: 09/838,858
;  PRIOR FILING DATE: 2001-04-20
;  PRIOR APPLICATION NUMBER: 09/756,096
;  PRIOR FILING DATE: 2001-01-08
;  PRIOR APPLICATION NUMBER: 09/158,863
;  PRIOR FILING DATE: 1998-09-23
;  PRIOR APPLICATION NUMBER: 09/133,717
;  PRIOR FILING DATE: 1998-08-13
;  PRIOR APPLICATION NUMBER: 09/087,233
;  PRIOR FILING DATE: 1998-05-28
;  PRIOR APPLICATION NUMBER: 08/766,354
;  PRIOR FILING DATE: 1996-12-13
;  NUMBER OF SEQ ID NOS: 125
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 119
;   LENGTH: 104
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Binding domain of human papilloma virus PTM
US-09-941-492-119
```

Query Match 76.0%; Score 15.2; DB 11; Length 104;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db         25 TGTCAAAAGCCACTGTGTCC 6
```

RESULT 14
US-09-941-492-122/c
; Sequence 122, Application US/09941492
; Publication No. US20030027250A1
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Lloyd
; APPLICANT: Garcia-Blanco, Mariano M.
; APPLICANT: Puttaraju, Madaiah
; APPLICANT: Mansfield, Gary S.
; TITLE OF INVENTION: METHODS OF COMPOSITIONS FOR USE IN
; TITLE OF INVENTION: SPLICEOSOME MEDIATED RNA TRANS-SPLICING
; FILE REFERENCE: A31304-BAE (072874.0156)
; CURRENT APPLICATION NUMBER: US/09/941,492
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/838,858
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/756,096
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 09/158,863
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 09/133,717
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 09/087,233
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 08/766,354
; PRIOR FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 122
; LENGTH: 138
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Binding domain of human papilloma virus PTM
US-09-941-492-122

Query Match 76.0%; Score 15.2; DB 11; Length 138;
Best Local Similarity 85.0%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db         60 TGTCAAAAGCCACTGTGTCC 41
```

RESULT 15
US-09-918-995-33830/c
; Sequence 33830, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 33830
; LENGTH: 431
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-33830

Query Match 76.0%; Score 15.2; DB 11; Length 431;
Best Local Similarity 85.0%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||||  || ||||||
Db        297 TGTCAAAAAGAGTGGTGTCC 278
```

Search completed: August 23, 2003, 13:53:16
Job time : 147 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: August 23, 2003, 10:18:48 ; Search time 50 Seconds
(without alignments)
176.553 Million cell updates/sec

Title: US-09-927-585A-1
Perfect score: 20
Sequence: 1 tgtcaaaaaccgttgtgtcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	29	1	US-08-495-743-27	Sequence 27, Appl
c	2	15.2	76.0	29	1	US-08-495-739-27	Sequence 27, Appl
c	3	15.2	76.0	29	1	US-08-495-741-27	Sequence 27, Appl
c	4	15.2	76.0	29	3	US-08-062-023-27	Sequence 27, Appl
	5	15.2	76.0	456	3	US-09-056-105-29	Sequence 29, Appl
	6	15.2	76.0	456	4	US-09-430-201-6	Sequence 6, Appli
	7	15.2	76.0	519	3	US-08-860-165-11	Sequence 11, Appl
	8	15.2	76.0	519	4	US-09-359-382-11	Sequence 11, Appl
	9	15.2	76.0	525	4	US-09-669-751-71	Sequence 71, Appl
	10	15.2	76.0	570	5	PCT-US94-05085A-1	Sequence 1, Appli
	11	15.2	76.0	570	5	PCT-US94-05085-1	Sequence 1, Appli
	12	15.2	76.0	776	1	US-08-216-233C-1	Sequence 1, Appli
	13	15.2	76.0	779	4	US-09-210-168-3	Sequence 3, Appli
	14	15.2	76.0	790	1	US-08-117-083-7	Sequence 7, Appli
	15	15.2	76.0	801	3	US-08-860-165-9	Sequence 9, Appli
	16	15.2	76.0	801	4	US-09-359-382-9	Sequence 9, Appli
	17	15.2	76.0	822	4	US-09-485-885-3	Sequence 3, Appli
	18	15.2	76.0	879	4	US-09-485-885-9	Sequence 9, Appli
	19	15.2	76.0	1000	1	US-07-965-274-3	Sequence 3, Appli
	20	15.2	76.0	1007	1	US-08-663-871-1	Sequence 1, Appli
	21	15.2	76.0	1007	1	US-08-664-894-1	Sequence 1, Appli
	22	15.2	76.0	1116	4	US-09-485-885-5	Sequence 5, Appli
	23	15.2	76.0	1173	4	US-09-485-885-13	Sequence 13, Appl
c	24	15.2	76.0	2368	2	US-08-520-933-1	Sequence 1, Appli
c	25	15.2	76.0	2368	4	US-09-285-040-1	Sequence 1, Appli
c	26	15.2	76.0	2368	6	5262177-1	Patent No. 5262177
c	27	15.2	76.0	2550	1	US-08-188-228-53	Sequence 53, Appl
c	28	15.2	76.0	2550	1	US-08-332-643-47	Sequence 47, Appl
c	29	15.2	76.0	2550	1	US-08-332-638-53	Sequence 53, Appl
	30	15.2	76.0	2832	3	US-08-819-177-2	Sequence 2, Appli
c	31	15.2	76.0	3136	1	US-08-188-228-41	Sequence 41, Appl
c	32	15.2	76.0	3136	1	US-08-332-638-41	Sequence 41, Appl
	33	15.2	76.0	7833	1	US-08-074-879-9	Sequence 9, Appli
	34	15.2	76.0	7833	1	US-08-468-057A-9	Sequence 9, Appli
	35	15.2	76.0	7904	1	US-08-316-239B-1	Sequence 1, Appli
	36	15.2	76.0	7904	1	US-08-316-239B-2	Sequence 2, Appli
	37	15.2	76.0	7904	1	US-08-410-005-1	Sequence 1, Appli
	38	15.2	76.0	7904	3	US-08-929-140-1	Sequence 1, Appli
	39	15.2	76.0	7904	3	US-09-560-579A-1	Sequence 1, Appli
	40	15.2	76.0	10803	3	US-09-080-044-1	Sequence 1, Appli
	41	15.2	76.0	10803	4	US-09-531-857A-1	Sequence 1, Appli
	42	14.8	74.0	1028	3	US-08-990-823-69	Sequence 69, Appl
	43	14.8	74.0	1028	4	US-09-477-135A-69	Sequence 69, Appl
	44	14.8	74.0	1107	4	US-09-000-094-19	Sequence 19, Appl
	45	14.8	74.0	1128	4	US-09-000-094-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-495-743-27/c
; Sequence 27, Application US/08495743
; Patent No. 5654416
; GENERAL INFORMATION:
; APPLICANT: Thomas J. Cummins, Susan M. Atwood
; APPLICANT: Lynn Bergmeyer, John B. Findlay
; APPLICANT: John W.H. Sutherland, JoAnne H. Kerschner
; TITLE OF INVENTION: DIAGNOSTIC COMPOSITIONS, ELEMENTS,
; TITLE OF INVENTION: METHODS AND TEST KITS FOR
; TITLE OF INVENTION: AMPLIFICATION AND DETECTION OF TWO
; TITLE OF INVENTION: OR MORE TARGET DNA'S USING PRIMERS
; TITLE OF INVENTION: HAVING MATCHED MELTING TEMPERATURES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eastman Kodak Company, Patent Legal Staff
; STREET: 343 State Street
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14650 - 2201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5inch, 1.44 MB storage (IBM)
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS Version 3.3
; SOFTWARE: PC-8 (Word for Windows)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/495,743
; FILING DATE: 08-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/062,023
; FILING DATE: 14-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Tucker, J. Lanny
; REGISTRATION NUMBER: 27,678
; REFERENCE/DOCKET NUMBER: 67271A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 722-9332
; TELEFAX: (716) 477-4646
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 nucleotides
; TYPE: Nucleic acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; MOLECULE TYPE: Primer for hPV DNA
; HYPOTHETICAL: No
; ANTI-SENSE: No
; ORIGINAL SOURCE: Synthetically prepared

```
;    IMMEDIATE SOURCE:  Same
;    PUBLICATION INFORMATION:  No. 5654416e
US-08-495-743-27

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 29;
  Best Local Similarity   85.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db         20 TGTCAAAAGCCACTGTGTCC 1


RESULT 2
US-08-495-739-27/c
; Sequence 27, Application US/08495739
; Patent No. 5702901
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas J. Cummins, Susan M. Atwood
;    APPLICANT:  Lynn Bergmeyer, John B. Findlay
;    APPLICANT:  John W.H. Sutherland, JoAnne H. Kerschner
;    TITLE OF INVENTION:  DIAGNOSTIC COMPOSITIONS, ELEMENTS,
;    TITLE OF INVENTION:  METHODS AND TEST KITS FOR
;    TITLE OF INVENTION:  AMPLIFICATION AND DETECTION OF TWO
;    TITLE OF INVENTION:  OR MORE TARGET DNA'S USING PRIMERS
;    TITLE OF INVENTION:  HAVING MATCHED MELTING TEMPERATURES
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eastman Kodak Company, Patent Legal Staff
;      STREET:  343 State Street
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  14650 - 2201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5inch, 1.44 MB storage (IBM)
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS Version 3.3
;      SOFTWARE:  PC-8 (Word for Windows)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/495,739
;      FILING DATE:  08-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/062,023
;      FILING DATE:  14 MAY 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tucker, J. Lanny
;      REGISTRATION NUMBER:  27,678
;      REFERENCE/DOCKET NUMBER:  67271A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 722-9332
;      TELEFAX:  (716) 477-4646
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 nucleotides
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Primer for hPV DNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:  Synthetically prepared
;    IMMEDIATE SOURCE:  Same
;    PUBLICATION INFORMATION:  No. 5702901e
US-08-495-739-27

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 29;
  Best Local Similarity   85.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db         20 TGTCAAAAGCCACTGTGTCC 1


RESULT 3
US-08-495-741-27/c
; Sequence 27, Application US/08495741
; Patent No. 5733751
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas J. Cummins, Susan M. Atwood
;    APPLICANT:  Lynn Bergmeyer, John B. Findlay
;    APPLICANT:  John W.H. Sutherland, JoAnne H. Kerschner
;    TITLE OF INVENTION:  DIAGNOSTIC COMPOSITIONS, ELEMENTS,
;    TITLE OF INVENTION:  METHODS AND TEST KITS FOR
;    TITLE OF INVENTION:  AMPLIFICATION AND DETECTION OF TWO
;    TITLE OF INVENTION:  OR MORE TARGET DNA'S USING PRIMERS
;    TITLE OF INVENTION:  HAVING MATCHED MELTING TEMPERATURES
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eastman Kodak Company, Patent Legal Staff
;      STREET:  343 State Street
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  14650 - 2201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5inch, 1.44 MB storage (IBM)
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS Version 3.3
;      SOFTWARE:  PC-8 (Word for Windows)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/495,741
;      FILING DATE:  08-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/062,023
;      FILING DATE:  19930514
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tucker, J. Lanny
;      REGISTRATION NUMBER:  27,678
;      REFERENCE/DOCKET NUMBER:  67271A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 722-9332
;      TELEFAX:  (716) 477-4646
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 nucleotides
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Primer for hPV DNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:  Synthetically prepared
;    IMMEDIATE SOURCE:  Same
;    PUBLICATION INFORMATION:  No. 5733751e
US-08-495-741-27

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 29;
  Best Local Similarity   85.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db         20 TGTCAAAAGCCACTGTGTCC 1


RESULT 4
US-08-062-023-27/c
; Sequence 27, Application US/08062023
; Patent No. 6174668
;  GENERAL INFORMATION:
;    APPLICANT:  Thomas J. Cummins, Susan M. Atwood
```

```
;    APPLICANT:  Lynn Bergmeyer, John B. Findlay
;    APPLICANT:  John W.H. Sutherland, JoAnne H. Kerschner
;    TITLE OF INVENTION:  DIAGNOSTIC COMPOSITIONS, ELEMENTS,
;    TITLE OF INVENTION:  METHODS AND TEST KITS FOR
;    TITLE OF INVENTION:  AMPLIFICATION AND DETECTION OF TWO
;    TITLE OF INVENTION:  OR MORE TARGET DNA'S USING PRIMERS
;    TITLE OF INVENTION:  HAVING MATCHED MELTING TEMPERATURES
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eastman Kodak Company, Patent Legal Staff
;      STREET:  343 State Street
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  14650 - 2201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5inch, 1.44 MB storage (IBM)
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS Version 3.3
;      SOFTWARE:  PC-8 (Word for Windows)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/062,023
;      FILING DATE:  19930514
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Tucker, J. Lanny
;      REGISTRATION NUMBER:  27,678
;      REFERENCE/DOCKET NUMBER:  67271A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 722-9332
;      TELEFAX:  (716) 477-4646
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 nucleotides
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  Primer for hPV DNA
;    HYPOTHETICAL:  No
;    ANTI-SENSE:  No
;    ORIGINAL SOURCE:  Synthetically prepared
;    IMMEDIATE SOURCE:  Same
;    PUBLICATION INFORMATION:  No. 6174668e
US-08-062-023-27

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 29;
  Best Local Similarity   85.0%;  Pred. No. 36;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db         20 TGTCAAAAGCCACTGTGTCC 1


RESULT 5
US-09-056-105-29
; Sequence 29, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
;  APPLICANT: KIPPS, THOMAS J.
;  APPLICANT: WU, YUNQI
;  TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
;  TITLE OF INVENTION: PROCESSING
;  FILE REFERENCE: 233/221
;  CURRENT APPLICATION NUMBER: US/09/056,105
;  CURRENT FILING DATE: 1998-04-06
;  EARLIER APPLICATION NUMBER: 60/043,467
;  EARLIER FILING DATE: 1997-04-10
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-056-105-29

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 456;
  Best Local Similarity   85.0%;  Pred. No. 51;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db        316 TGTCAAAAGCCACTGTGTCC 335


RESULT 6
US-09-430-201-6
; Sequence 6, Application US/09430201
; Patent No. 6313373
; GENERAL INFORMATION:
;  APPLICANT:  Eckert, Richard L.
;  APPLICANT:  Crish, James F.
;  TITLE OF INVENTION:  Tissue Specific Promoters and Transgenic Animals for
;  TITLE OF INVENTION:  the Screening of Pharmaceuticals
;  FILE REFERENCE:  CASE-04022
;  CURRENT APPLICATION NUMBER: US/09/430,201
;  CURRENT FILING DATE:  1999-10-29
;  PRIOR APPLICATION NUMBER:  60/106,495
;  PRIOR FILING DATE:  1998-10-30
;  NUMBER OF SEQ ID NOS:  8
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 6
;   LENGTH: 456
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-430-201-6

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 456;
  Best Local Similarity   85.0%;  Pred. No. 51;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              ||||||||  ||  |||||||
Db        316 TGTCAAAAGCCACTGTGTCC 335


RESULT 7
US-08-860-165-11
; Sequence 11, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
;  APPLICANT: EDWARDS, Stirling John
;  APPLICANT: COX, John Cooper
;  APPLICANT: WEBB, Elizabeth Ann
;  APPLICANT: FRAZER, Ian
;  TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
;  FILE REFERENCE: 17227/130
;  CURRENT APPLICATION NUMBER: US/08/860,165A
;  CURRENT FILING DATE: 1997-09-22
;  EARLIER APPLICATION NUMBER: PCT/AU95/00868
;  EARLIER FILING DATE: 1995-12-20
;  EARLIER APPLICATION NUMBER: AU PN0157
;  EARLIER FILING DATE: 1994-12-20
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(516)
;   FEATURE:
```

```
;   OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-11

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 519;
  Best Local Similarity   85.0%;  Pred. No. 52;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        151 TGTCAAAAGCCACTGTGTCC 170


RESULT 8
US-09-359-382-11
; Sequence 11, Application US/09359382
; Patent No. 6306397
; GENERAL INFORMATION:
;  APPLICANT: EDWARDS, Stirling John
;  APPLICANT: COX, John Cooper
;  APPLICANT: WEBB, Elizabeth Ann
;  APPLICANT: FRAZER, Ian
;  TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
;  FILE REFERENCE: 017227/0148
;  CURRENT APPLICATION NUMBER: US/09/359,382
;  CURRENT FILING DATE: 1999-07-23
;  EARLIER APPLICATION NUMBER: US 08/860,165
;  EARLIER FILING DATE: 1997-09-22
;  EARLIER APPLICATION NUMBER: PCT/AU95/00868
;  EARLIER FILING DATE: 1995-12-20
;  EARLIER APPLICATION NUMBER: AU PN0157/94
;  EARLIER FILING DATE: 1994-12-20
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human papillomavirus type 16
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(516)
US-09-359-382-11

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 519;
  Best Local Similarity   85.0%;  Pred. No. 52;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        151 TGTCAAAAGCCACTGTGTCC 170


RESULT 9
US-09-669-751-71
; Sequence 71, Application US/09669751
; Patent No. 6551575
; GENERAL INFORMATION:
;  APPLICANT:  Greenspan, Ralph J.
;  TITLE OF INVENTION:  Methods for Identifying Compounds for
;  TITLE OF INVENTION:  Motion Sickness, Vertigo and Other Disorders Related to
;  TITLE OF INVENTION:  Balance and the Perception of Gravity
;  FILE REFERENCE:  P-NI 3864
;  CURRENT APPLICATION NUMBER: US/09/669,751
;  CURRENT FILING DATE:  2000-09-26
;  PRIOR APPLICATION NUMBER:  US 60/168,579
;  PRIOR FILING DATE:  1999-12-02
;  NUMBER OF SEQ ID NOS:  261
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 71
;   LENGTH: 525
;   TYPE: DNA
;   ORGANISM: Drosophila
US-09-669-751-71

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 525;
  Best Local Similarity   85.0%;  Pred. No. 52;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              | |||||||| ||||||| |
Db        244 TCTCAAAAACGGTTGTGTTC 263


RESULT 10
PCT-US94-05085A-1
; Sequence 1, Application PC/TUS9405085A
;  GENERAL INFORMATION:
;    APPLICANT:  Janice T. Brown
;    TITLE OF INVENTION:  HUMAN PAPILLOMAVIRUS DETECTION ASSAY
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Baxter Diagnostics Inc.
;      STREET:  One Baxter Parkway, Building DP-3E
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Apple Macintosh System 7.0
;      SOFTWARE:  Macintosh Text File
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05085A
;      FILING DATE:  N/A
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/058,920
;      FILING DATE:  May 6, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mark Buonaiuto
;      REGISTRATION NUMBER:  31,593
;      REFERENCE/DOCKET NUMBER:  BA-4448
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/948-2537
;      TELEFAX:  708/948-2642
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  570
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  Papaoviridae, Human papilloma virus
;      STRAIN:  16
;    FEATURE:
;      NAME/KEY:  Portion of viral genome coding for E6/E7 polypeptides.
;    PUBLICATION INFORMATION:
;      AUTHORS:  Seedorf, K., Krammer, G., Durst, M.,
;      AUTHORS:  Suhai, S., and Rowekamp, W.
;      TITLE:  Human Papillomavirus Type 16 DNA Sequence
;      JOURNAL:  Virology
;      VOLUME:  145
;      ISSUE:
;      PAGES:  181-185
;      DATE:  1985
PCT-US94-05085A-1

  Query Match             76.0%;  Score 15.2;  DB 5;  Length 570;
  Best Local Similarity   85.0%;  Pred. No. 53;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        338 TGTCAAAAGCCACTGTGTCC 357
```

```
RESULT 11
PCT-US94-05085-1
; Sequence 1, Application PC/TUS9405085
;  GENERAL INFORMATION:
;    APPLICANT:  Janice T. Brown
;    TITLE OF INVENTION:  HUMAN PAPILLOMAVIRUS DETECTION ASSAY
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Baxter Diagnostics Inc.
;      STREET:  One Baxter Parkway, Building DP-3E
;      CITY:  Deerfield
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  Apple Macintosh System 7.0
;      SOFTWARE:  Macintosh Text File
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05085
;      FILING DATE:  06-MAY-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/058,920
;      FILING DATE:  May 6, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mark Buonaiuto
;      REGISTRATION NUMBER:  31,593
;      REFERENCE/DOCKET NUMBER:  BA-4448
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  708/948-2537
;      TELEFAX:  708/948-2642
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  570
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    HYPOTHETICAL:  no
;    ANTI-SENSE:  no
;    ORIGINAL SOURCE:
;      ORGANISM:  Papaoviridae, Human papilloma virus
;      STRAIN:  16
;    FEATURE:
;      NAME/KEY:  Portion of viral genome coding for E6/E7 polypeptides.
;    PUBLICATION INFORMATION:
;      AUTHORS:  Seedorf, K., Krammer, G., Durst, M.,
;      AUTHORS:  Suhai, S., and Rowekamp, W.
;      TITLE:  Human Papillomavirus Type 16 DNA Sequence
;      JOURNAL:  Virology
;      VOLUME:  145
;      ISSUE:
;      PAGES:  181-185
;      DATE:  1985
PCT-US94-05085-1

  Query Match             76.0%;  Score 15.2;  DB 5;  Length 570;
  Best Local Similarity   85.0%;  Pred. No. 53;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        338 TGTCAAAAGCCACTGTGTCC 357


RESULT 12
US-08-216-233C-1
; Sequence 1, Application US/08216233C
; Patent No. 5506105
;  GENERAL INFORMATION:
;    APPLICANT:  HAYDOCK, PAUL V.
;    TITLE OF INVENTION:  IN SITU ASSAY OF AMPLIFIED INTRACELLULAR
;    TITLE OF INVENTION:  mRNA TARGETS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DADE INTERNATIONAL INC.
;      STREET:  1717 DEERFIELD ROAD
;      CITY:  DEERFIELD
;      STATE:  ILLINOIS
;      COUNTRY:  USA
;      ZIP:  60015
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/216,233C
;      FILING DATE:  22-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/808,456
;      FILING DATE:  10-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TYMESON, CYNTHIA G.
;      REGISTRATION NUMBER:  34,745
;      REFERENCE/DOCKET NUMBER:  BA-4203
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (305) 222-6423
;      TELEFAX:  (305) 222-6686
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  776 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-216-233C-1

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 776;
  Best Local Similarity   85.0%;  Pred. No. 55;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        337 TGTCAAAAGCCACTGTGTCC 356


RESULT 13
US-09-210-168-3
; Sequence 3, Application US/09210168
; Patent No. 6355424
; GENERAL INFORMATION:
;  APPLICANT:  Lorincz, Attila T.
;  TITLE OF INVENTION:  ASSESSMENT OF HUMAN PAPILLOMA VIRUS-RELATED DISEASE
;  FILE REFERENCE:  26294005US2
;  CURRENT APPLICATION NUMBER: US/09/210,168
;  CURRENT FILING DATE:  1998-12-11
;  PRIOR APPLICATION NUMBER:  U.S. 60/082,167
;  PRIOR FILING DATE:  1998-04-17
;  PRIOR APPLICATION NUMBER:  U.S. 60/070,486
;  PRIOR FILING DATE:  1998-01-05
;  PRIOR APPLICATION NUMBER:  U.S. 60/069,426
;  PRIOR FILING DATE:  1997-12-12
;  NUMBER OF SEQ ID NOS:  5
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 779
;   TYPE: DNA
;   ORGANISM: Human papillomavirus
;   FEATURE:
;   OTHER INFORMATION: E6/E7-HPV16
US-09-210-168-3
```

```
  Query Match             76.0%;  Score 15.2;  DB 4;  Length 779;
  Best Local Similarity   85.0%;  Pred. No. 55;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        344 TGTCAAAAGCCACTGTGTCC 363


RESULT 14
US-08-117-083-7
; Sequence 7, Application US/08117083
; Patent No. 5719054
;  GENERAL INFORMATION:
;    APPLICANT:  Boursnell, Michael E.
;    APPLICANT:  Inglis, Stephen C.
;    APPLICANT:  Munro, Alan J.
;    TITLE OF INVENTION:  Recombinant Virus Vectors Encoding Human
;    TITLE OF INVENTION:  Papilloma Virus Proteins
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Walter H. Dreger
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/117,083
;      FILING DATE:  10-SEP-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Walter H.
;      REGISTRATION NUMBER:  24,190
;      REFERENCE/DOCKET NUMBER:  A-58783
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  790 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-117-083-7

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 790;
  Best Local Similarity   85.0%;  Pred. No. 55;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        342 TGTCAAAAGCCACTGTGTCC 361


RESULT 15
US-08-860-165-9
; Sequence 9, Application US/08860165A
; Patent No. 6004557
; GENERAL INFORMATION:
;  APPLICANT: EDWARDS, Stirling John
;  APPLICANT: COX, John Cooper
;  APPLICANT: WEBB, Elizabeth Ann
;  APPLICANT: FRAZER, Ian
;  TITLE OF INVENTION: VARIANTS OF HUMAN PAPILLOMA VIRUS ANTIGENS
;  FILE REFERENCE: 17227/130
;  CURRENT APPLICATION NUMBER: US/08/860,165A
;  CURRENT FILING DATE: 1997-09-22
;  EARLIER APPLICATION NUMBER: PCT/AU95/00868
;  EARLIER FILING DATE: 1995-12-20
;  EARLIER APPLICATION NUMBER: AU PN0157
;  EARLIER FILING DATE: 1994-12-20
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 801
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(798)
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Gene Fusion
US-08-860-165-9

  Query Match             76.0%;  Score 15.2;  DB 3;  Length 801;
  Best Local Similarity   85.0%;  Pred. No. 55;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 TGTCAAAAACCGTTGTGTCC 20
              |||||||| ||  |||||||
Db        337 TGTCAAAAGCCACTGTGTCC 356


Search completed: August 23, 2003, 12:15:08
Job time : 51 secs
```